Who Debates, Who Wins? At-Scale Experimental Evidence on the Supply of Policy Information in a Liberian Election*

Jeremy Bowles† & Horacio Larreguy‡

October 2, 2020

Abstract

We examine how candidate selection into the supply of policy information determines its electoral effects. In a nationwide debate initiative designed to solicit and rebroadcast policy promises from Liberian legislative candidates, we randomized the encouragement of debate participation across districts. The intervention substantially increased the debate participation of leading candidates but led to uneven electoral returns for these candidates, with incumbents benefiting at the expense of challengers. These results are driven by differences in compliance: complying incumbents, but not challengers, positively selected into debate participation based on the alignment of their policy priorities with those of their constituents.

JEL codes: D72, O12

Keywords: accountability, information, selection

*We are grateful to the J-PAL Governance Initiative and USAID for funding the data collection. Larreguy gratefully acknowledges funding from the French Agence Nationale de la Recherche under the Investissement d'Avenir program ANR-17-EURE-0010. With thanks to GeoPoll, Internews Liberia, the National Elections Commission of Liberia, NORC at the University of Chicago, the USAID Democracy, Human Rights and Governance Office and USAID Liberia for their cooperation. We thank Kate Casey, Cesi Cruz, Patrick Francois, Thomas Fujiwara, Miriam Golden, Torben Iversen, Eric Kramon, Julien Labonne, Evan Lieberman, John Marshall, Vincent Pons, Pablo Querubin, Pia Raffler, Franceso Trebbi and Ariel White for useful discussions, and audiences at APSA, the Harvard Experiments Conference, the Harvard Political Economy of Development Working Group, Internews, Liberia Strategic Analysis, the NYU Center for Experimental Social Science, Toulouse School of Economics, Twaweza, USAID Liberia, University of British Columbia, Universitat Autonoma de Barcelona, Universitat Pompeu Fabra, the Wallis Institute Annual Conference and WGAPE at NYU-Abu Dhabi for their feedback. With special thanks to Morgan Holmes and Jan McArthur for their enduring support, to Anthony Chan, Michael Haines, Anna Langer, Tawedzegwa Musitini and April O'Neill for their cooperation and to Luis da Silva, Theo Harris, Alpha Senkpeni, Kate Thomas, and Jesús Ynfusino for research assistance. IRB approval granted by Harvard Committee on the Use of Human Subjects (ID: IRB17-1178) and NORC at the University of Chicago (ID: 7554.072.01). Study pre-registered with EGAP (ID: 20171024AA) and AEA (ID: AEARCTR-0002553).

†PhD Candidate, Department of Government, Harvard University. jbowles@g.harvard.edu.

‡Associate Professor, Department of Government, Harvard University, and Visiting Researcher, Institute for Advanced Study in Toulouse and Toulouse School of Economics. hlarreguy@fas.harvard.edu and hlarreguy@tse-fr.eu.

1 Introduction

Following classic models of electoral accountability (Barro, 1973; Ferejohn, 1986; Fearon, 1999; Holmstrom, 1999), much extant empirical research considers how supplying citizens with policy information, generally about their incumbents, affects what they know about candidates, how they evaluate those candidates, and ultimately their voting behavior.[1] While often demonstrated to promote programmatic competition, this work overlooks the strategic decision that candidates make in supplying policy information to the public in the first place. In advanced developed democracies, candidates may face no choice but to supply policy information. In more weakly-institutionalized settings, however, selection into the supply of policy information can neither be assumed nor enforced. Designing scalable initiatives that promote programmatic competition in such settings requires a better understanding of the incentives candidates face in choosing to supply policy information and, accordingly, the electoral consequences of doing so.

Candidate participation in such initiatives cannot be taken for granted. The decision to supply policy information is risky, particularly in clientelistic settings where the returns to programmatic competition might be limited and uncertain (Fujiwara and Wantchekon, 2013). Candidates *ex ante* risk performing poorly, revealing their policy priorities to be disconnected from their constituents', and restricting their ability to target campaign promises to small groups of influential voters. These risks are especially pronounced for leading candidates, who emerged within the existing clientelistic equilibrium, enjoy greater resources for on-the-ground campaigning, and are more likely to be attacked by opponents to gain publicity. Focusing on election debate initiatives, which have spread rapidly across developing democracies as a way to disseminate policy information to voters, accordingly, one leading debate organizer notes that "the greatest universal challenge that [debate] sponsors face regardless of country or culture is convincing candidates to take part" (National Democratic Institute, 2014). Illustratively, incumbent candidates in sub-Saharan Africa have refused to participate in nearly half of all presidential debates.

We study the electoral consequences of the supply of policy information by legislative candidates in Liberia, where clientelism is pervasive and the media sector remains weak. In an effort to enhance programmatic electoral competition in the 2017 legislative election, USAID

[1] Ferraz and Finan (2008), and Larreguy et al. (2019) show that local media reports on incumbent performance enhance electoral accountability. However, a series of localized information dissemination campaigns fail to consistently replicate these media effects (Banerjee et al., 2011; Chong et al., 2015; Bhandari et al., 2019; Dunning et al., 2019b). Closest to our work, recent studies analyze whether debates can lead citizens to make more informed voting decisions (Bidwell et al., 2019; Platas and Raffler, 2019).

and the international NGO Internews organized standardized debates in every electoral district. The moderators asked participating candidates a series of questions regarding their policy priorities, which often included district schools, primary healthcare facilities, and infrastructural investments. Rather than large townhall-style debates, the emphasis was on soliciting concrete policy promises from the candidates that would then be rebroadcast multiple times by community radio stations.

We partnered with USAID and Internews to experimentally study how the effects of their large-scale initiative were shaped by the participation of leading candidates, i.e. incumbents and and their most serious challengers, which they were concerned would be low. The participation of these candidates in the debates was considered important, first, so that citizens would pay attention to the debates and, likely, engage more with the debate information as a result. Second, because few other channels for disseminating policy information exist, debate participation was a primary way that constituents could learn about candidates' policy platforms, update their beliefs about different candidates, and then potentially vote for those candidates with priorities closest to their own.

To examine how candidate participation affected the electoral consequences of the debates, we randomly assigned an intervention designed to more intensively persuade candidates to participate in treated districts compared to normal intensity in inviting candidates in control districts. Our intervention was aimed at mitigating the concerns of leading candidates regarding personal attacks and biased moderators, and at highlighting that the emphasis of the debates would be narrowly focused on policy proposals. To evaluate the effects of the intervention, we make use of a rich set of original data sources, including polling station-level electoral results, a nationwide panel survey of more than 4,000 citizens, a survey of more than 600 candidates who ran in the election, a survey of more than 50 radio stations, full transcripts from debates, and around 20 focus groups.

We find that the intervention had striking, and uneven, electoral consequences on *leading candidates*.[2] The intervention increased the participation of incumbents by a third to 50%, while the participation of those who we predicted to be their most credible challengers increased by a fifth to 60%. There was a negligible treatment effect for marginal candidates, who participated at higher rates absent the intervention. Ultimately, the intervention led to improved electoral

[2]We provide a detailed discussion of why we focus on, and how we categorized, *leading candidates* in Section 2.1 and Appendix A.1. In short, for both theoretical and logistical reasons, we focus primarily on the predicted top three candidates per race. This comprised incumbent candidates ($n = 64$) and their predicted challengers ($n = 155$), which together we denote *leading candidates* ($n = 219$) and represent a subset of the total 984 candidates who sought office.

outcomes for incumbents in the treated districts and worsened outcomes for their challengers. Consistent across our polling station data and voter surveys, treated citizens voted for incumbents around 4 percentage points (pp) more often and for their challengers around 3 pp less often. These changes in electoral outcomes occurred particularly in districts where we predicted incumbents to perform well in the debates,[3] and where our survey evidence indicates that incumbents had policy platforms that were better aligned with the priorities of their constituents. Remarkably, 50% of incumbents in treated districts won re-election, compared to 43% in control districts.

We then establish three sets of results to explain these electoral effects, which contrast with the results of prior experimental interventions. First, we show differences in compliance with the intervention between incumbents and challengers. Incumbents who complied with the intervention were not well known by their constituents beforehand, but their policy priorities were well aligned with those of their constituents. In other words, incumbents recognized whether their policy priorities matched those of their constituents and whether their constituents knew this. In turn, complying challenger candidates, about whom citizens were also very uncertain, apparently lacked the political sophistication to correctly assess the returns to debate participation. If anything, they negatively self-selected into debate participation on the basis of the overlap between their policy priorities and those of their constituents.[4]

Second, we show that, likely as a result of increased participation by leading candidates, citizens in treated districts paid more attention to the debates and acquired more political information. Citizens in treated districts were 0.3 standard deviations (sd) more intensively exposed to debate content than those in control districts. Similarly, citizens in treated districts were 0.13 sd and 0.15 sd respectively more likely to exhibit knowledge about the debates themselves and of poorly-understood policy issues discussed in each debate than citizens in control districts. As our qualitative accounts corroborate, citizens were more interested in and responsive to debate information when leading candidates participated. In treated districts, citizens engaged in 0.25 sd greater political information acquisition. Consistent with prior studies of legislative debates, this increased exposure led citizens in treated districts to update their beliefs about candidates (Bidwell et al., 2019; Platas and Raffler, 2019). Citizens in treated districts updated their beliefs positively (but weakly) about incumbents' competence and policy

[3] We do this by running a LASSO model of the debate performance outcomes of participating candidates on pre-determined covariates.

[4] This pattern of compliance largely mirrors the self-selection of candidates into debate participation in control districts. This suggests that the treatment effects of the intervention are likely representative of those of the overall debate initiative.

priorities, becoming more certain in their knowledge in the process. However, in treated districts, citizens experienced negligible gains in certainty about the competence or policy priorities of challengers, and updated weakly *negatively* about their competence. Third, aided by an increase in demand from radio stations, incumbents increased their radio campaigning in treated districts while challengers reduced their on-the-ground campaigning. This suggests that incumbent debate participation had a deterrent effect on challengers' campaign efforts.

The evidence is consistent with the variation in candidates' ability to correctly evaluate the returns to participation critically determining the electoral consequences of the debate initiative. This differential selection into the supply of policy information may also condition the effect of similar large-scale initiatives, whether candidate debates or other initiatives designed to enhance programmatic political competition. Most of the non-participating incumbents had policy preferences that were very poorly aligned with those of their constituents. Had all candidates participated, therefore, the results could even have reversed what we find. Such results would match those of randomized small-scale debate interventions, all of which ensured that leading candidates participated (Bidwell et al., 2019; Platas and Raffler, 2019). Our research shows that candidate attendance, and therefore the external validity of such results, cannot be presumed when scaling such initiatives because their effects depend on the attendance of these leading candidates. Indeed, districts that were not assigned to the invitation intervention saw only 35% of incumbents and 50% of challengers participating in the debate.

The extent and implications of variation in candidates' selection into programmatic initiatives are likely to differ by context, and we do not seek to extrapolate the exact pattern of our results to other settings. However, our findings represent a proof of concept in the context of recent work highlighting the importance of experimentation at scale, since the effect of small-scale interventions might differ substantially when scaled (Al-Ubaydli et al., 2017; Banerjee et al., 2017; Muralidharan and Niehaus, 2017). In particular, Al-Ubaydli et al. (2017) point to individual non-adoption of the treatment as a key factor explaining the difference in outcomes between small-scale and large-scale interventions. This lack of compliance may be especially pronounced when candidates face incentives not to participate in initiatives designed to move away from the clientelistic equilibrium from which they often benefit. Our results highlight the importance of investigating who selects into programmatic initiatives to understand their effects when scaled.

More specifically, our findings contribute to the literature on information and electoral accountability. This literature shows that such accountability is enhanced through revelations of

past incumbent performance via broadcast and social media (Ferraz and Finan, 2008; Enríquez et al., 2020; Larreguy et al., 2019), but not necessarily via localized information dissemination campaigns (Banerjee et al., 2011; de Figueiredo et al., 2013; Chong et al., 2015; Bhandari et al., 2019; Dunning et al., 2019a). Closer to our work, recent studies that analyzed the effect of small-scale debate initiatives in which the participation of the leading candidates was ensured found a leveling of the playing field favoring challenger candidates (Bidwell et al., 2019; Platas and Raffler, 2019). Other work similarly shows that citizens exposed to the campaign promises of all candidates learn about candidate policy priorities and vote accordingly (Cruz et al., 2019). However, none of these studies address how the results of their information dissemination initiatives might be conditioned by the strategic decision of candidates to supply such information or not, which in turn may affect their external validity.

Lastly, our results also contribute to recent experimental work designed to reduce clientelistic campaign practices common in developing democracies (Finan and Schechter, 2012; Cruz et al., 2018; Duarte et al., 2019). Fujiwara and Wantchekon (2013) show that programmatic platforms transmitted through town hall meetings reduced vote buying and increased electoral support for the candidates involved. Vicente (2014), Hicken et al. (2018), Blattman et al. (2019), Bobonis et al. (2019), and Vasudevan (2019), however, show that interventions designed to combat vote-buying practices while facilitating competition on the basis of policy vary in their success. Our results underscore the importance of variation in the suitability for programmatic competition among candidates who became prominent under the existing clientelistic equilibrium. This suggests that recruiting new citizens to run for office and training them in the articulation of policy platforms could be important forces in shifting towards more programmatic political equilibria.

2 Background

Since its emergence from civil war in 2003, Liberia has held three presidential elections, three legislative elections, and two senatorial elections. We focus on the House of Representatives election of October 2017, in which each of 73 electoral districts elected a single representative for a six-year term in a first-past-the-post electoral system. Key responsibilities of House members include making laws, controlling access to development funds, and allocating and implementing public goods. Representatives are rewarded with an annual salary over $200,000 USD in a country with an annual per capita income of $900 (IREDD, 2016). The relatively low

barriers to candidacy combined with a fragmented and weak party environment mean that there are usually a large number of candidates for each office. In the 2017 election there were 984 candidates from 26 different political parties, with as many as 28 candidates in a district. As one editorial opined, "Rest assured that this is not a healthy expression of diverse opinions. Everyone wants a piece of the pie." (Glencorse and Yealue, 2017).

Once in office, legislator performance is varied, with some lawmakers having close to perfect attendance while others attending as few as 45% of legislative sessions (IREDD, 2016). Dissatisfaction with incumbent performance is widespread, with two-thirds of citizens mistrusting their representative (Afrobarometer, 2015) and more than half of citizens report seeing their representatives only at election time, once every six years (USAID, 2018). This dissatisfaction is not the result of citizens being unaware of who their legislator is—92% of our citizen survey correctly named their legislator—but partially owes to the lack of credible information about political activity as a result of an underdeveloped media sector and poor infrastructural conditions. While incumbents sought re-election in nearly 90% of districts, less than half of them were reelected.

2.1 Who runs for office?

We draw on an original survey of 612 candidates to provide some descriptive evidence of candidate characteristics.[5] Throughout this paper we distinguish between three *predicted leading candidates* per district and *other candidates*, and further divide *predicted leading candidates* into *incumbents* and *predicted challengers*. This assignment of 219 predicted leading candidates—64 incumbents and 155 predicted challengers—and 765 *other candidates* is described fully in Appendix A.1. The objective with this categorization was to facilitate measurement and analysis by identifying three candidates per district who had genuine chances of success—qualitatively, there exists a long tail of candidates who run primarily to enhance their profiles and secure post-electoral favors (Spatz and Thaler, 2018). Moreover, the definition of *actual leading candidates*, those whose vote share ranked in the top three of their district, might be endogenous to our intervention.

In Table 1 we provide descriptive statistics by candidate category. Candidates generally come from Liberia's elite, with more education than an average citizen, and are overwhelmingly male (84%). Incumbents are older and possess higher levels of education than challenger candidates. They are much more likely to possess prior experience in a non-elected government

[5]The 62% response rate is balanced across treatment groups both overall and within candidate categories.

1. Candidate Characteristics

	Age (1)	University educated (2)	Ran before (3)	Govt. job before (4)	NGO job before (5)	Advocacy experience (6)	Campaign expenditure (7)	Radio station (8)
Incumbent	55.8	0.68	1.00	0.48	0.35	0.87	$61,458	0.16
Challenger	48.9	0.64	0.43	0.30	0.38	0.88	$41,282	0.06
Other	47.7	0.53	0.22	0.32	0.42	0.85	$30,083	0.03

Table displays mean values of column variables for incumbents, challengers, and other candidates according to our survey. 'Age:' candidate age in years. 'University educated:' candidate has completed university. 'Ran before:' candidate ran for office at least once before. 'Govt. job before:' candidate has held non-elected government job before. 'NGO job before:' candidate has worked for an NGO before. 'Advocacy experience:' candidate reports having worked on an advocacy campaign before. 'Campaign expenditure:' self-reported campaign spending in USD. 'Radio station:' candidate either owns or manages a radio station.

job and are less likely to have experience working for an NGO. Almost a third of all candidates have previously run for office, and most report experience in advocacy campaigns in their districts. Candidates report spending substantial amounts—on average above $30,000—on their campaigns. Incumbents, however, report spending 50% more than predicted challengers and 100% more than other candidates. The differences between incumbents and challengers are consistent with the substantial literature on the resource advantages enjoyed by incumbents in developing democracies.

2.2 Campaigning and policy promises

Legislative campaigns are primarily organized by the candidates themselves with little coordination and oversight by the relatively weak national-level parties. Electoral campaigns are marked by local rallies at which candidates distribute gifts in cash or kind to generate support. Nearly 80% of surveyed candidates reported visiting most or all communities in their district, while nearly half reported distributing gifts in most or all communities. During campaign season, incumbents in particular orchestrate mass campaigns to buy turnout and truck voters from the capital to their districts (Bowles et al., 2019). A USAID survey in 2015 found that 35% of respondents were personally given money in exchange for their vote (USAID, 2015).

In this clientelistic context, candidates face few incentives to widely disseminate policy promises. The candidates exhibit great awareness of the differences in the *types* and *credibility* of policy promises delivered at local rallies versus over the radio, as Table 2 shows based on responses to our survey of candidates. Overall, candidates believe that promises made on the radio are more credible than those made at rallies, but they acknowledge the low likelihood of any campaign promise being kept. Relevant for our later results, incumbents appear to be more

sophisticated in this regard.

2. Candidate Attitudes Towards Policy Promises

Candidate type	Different promises (1)	Rally credibility (2)	Radio credibility (3)
Incumbent	0.73	0.19	0.26
Challenger	0.70	0.12	0.14
Other	0.67	0.16	0.15

Table displays mean values of binary column variables for incumbent, challenger, and other candidate respondents based on our surveys. 'Different promises:' candidate believes that candidates make different promises on radio versus in-person campaigning. 'Rally credibility:' candidate believes that promises made by candidates at rallies are very likely to be fulfilled. 'Radio credibility:' candidate believes that promises made by candidates on radio are very likely to be fulfilled.

Because candidates lack incentives to publicize policy promises, the wide dissemination of policy platforms across districts is extremely rare.[6] As such, candidate campaigns generally lack policy platforms, focusing instead on promises of local development (which rarely come to fruition) delivered through on-the-ground campaigning buoyed by vote buying. The absence of programmatic information is facilitated by a fractured media landscape. Radio stations are a potentially important source of access to political information: radio ownership is high at 83%, and 62% of Liberian respondents report listening to news on the radio every day (Afrobarometer, 2015). However, because the radio industry lacks regulation, the market is fragmented, access to electricity is sporadic, and sources of commercial revenue are scarce, radio stations frequently become the mouthpieces of particular political figures and local firms (Kamara, 2017). Indeed, as Table 1 shows, many incumbents actually own their own radio stations and many more candidates are informally connected to other stations.

3 Candidate debates

In the run-up to the October 2017 elections for the House of Representatives, Internews Liberia led a nationwide debate initiative that they hoped would push back against Liberia's clientelistic equilibrium and establish the beginnings of a programmatic one by increasing the supply of policy information to citizens.[7] Internews partnered with several Liberian journalist associations

[6]One of the country's most prominent newspapers, the *Daily Observer*, built a "promises tracker" ahead of the election where candidates could outline their policy platforms. No incumbent did this.

[7]Other organizations held debates for the presidential race; those debates are not the focus here.

Figure 1. District Debate (Montserrado D3)

to organize debates across different parts of the country.[8] In each district, a local journalist was responsible for conducting research about the issues relevant to constituents, publicizing the debate, and moderating it. The debates took place from mid-August to mid-September 2017. In total, 129 debates were held across all 73 districts.[9]

Debate venues were mostly administrative buildings, town halls, and schools. Every debate followed a simple and uniform structure. First, candidates were asked to outline their campaign policy promises. The moderator then posed the same questions to each candidate in turn, and each candidate was allowed three minutes to respond. The first question in each debate was related to the management of the County Social Development Fund (CSDF), which is poorly managed, with little oversight or input from citizens. Second, candidates were asked about how they would spend their Legislative Support Project (LSP) discretionary funds. After these standardized questions, candidates were asked two or three questions about relevant local issues based on research conducted by the moderator. Moderators intervened to prevent candidates from making personal attacks on other candidates.

Debate content was disseminated by partnering with at least one community radio station in each district that would broadcast and later rebroadcast the debate on average six times,

[8]These partners were the Press Union of Liberia (PUL), Liberia Media for Democratic Initiatives (LMDI), and the Center for Media Studies and Peacebuilding (CEMESP).

[9]In districts with a large number of candidates, multiple debates were held, generally on the same day, with candidates randomly assigned to a debate.

with rebroadcasting concentrated in the 10 days before the election.[10] Internews chose 43 radio stations to rebroadcast the debates, making selections based on the signal strength of the station to maximize audience size and discarding any stations that were managed or owned by candidates running for office in that district.[11] We estimate that nearly 90% of the population was covered by a signal from the station broadcasting that district's debate.

3.1 Invitation intervention

Whether to participate in a debate represents a strategic decision by candidates based on whether disseminating policy information will help or hurt their electoral bid. In the context of Liberia, as in many developing democracies, leading candidates generally emerge by developing clientelistic machines that mobilize voters on their behalf (Bowles et al., 2019). The returns to programmatic competition for these candidates were deeply uncertain and, potentially, negative (Fujiwara and Wantchekon, 2013). Many leading candidates expressed hesitation regarding debate participation, and the partners were concerned that the attendance of the leading candidates at the debates would be low, which could undermine the potential impact of the initiative.

The risks to debate participation were particularly acute for the leading candidates, who were frequently concerned that their proposed policies and competence would be attacked by other candidates and that moderators would be biased against them. Furthermore, poor performance in the debate might hurt their candidacy, while committing to a policy platform would restrict their ability to deviate on both the campaign trail and, eventually, once in office. Participation also involved non-trivial direct costs given the difficulties of travel in Liberia's rainy season. Leading candidates, who possessed greater resources available for on-the-ground campaigning than other candidates, additionally faced greater opportunity costs of participation. For less-relevant candidates with more limited resources, debate participation offered a much clearer positive expected return: they lacked the resources to buy votes or hold rallies, so participating in the debate represented a significant opportunity for them to publicize their name and policy priorities.

Therefore, to evaluate the impact of leading candidates' selection into supplying programmatic information, we randomized the level of effort associated with debate invitation across

[10]The audience in attendance at each debate was around 100 people. Election-related violence is a concern in Liberia and so, to minimize the risk of conflict, in-person audiences for the debates were purposefully kept small.

[11]The debates were broadcast by fewer than 73 radio stations since some had the ability to broadcast debates in more than one district. A few stations were discarded due to political affiliations, primarily because they could not be relied upon to replay the debates in full with no editing.

districts.[12]

Rather than experimentally varying the extensive margin of whether candidates were invited at all, in our treated districts we significantly increased the intensity of invitation activities.

Candidates in control districts were contacted by the relevant Liberian journalist association who invited them to attend and provided logistical information about the debate. In treatment districts, this was augmented with three additional components aimed at persuading leading candidates to attend. First, we sent official invitations via email from Internews with USAID branding as far ahead of the debates as possible. These invitations included logistical details and instructions for candidates to contact Internews if they had any doubts about the debates.[13] Second, we made phone calls to all candidates around two days before each debate to persuade them to attend. These were mostly conducted by a high-profile Liberian radio journalist who is widely known and respected by politicians. These calls were designed to address any concerns candidates had about the debates and to clarify the objective, structure, and unbiasedness of the debates.[14] Third, we sent SMS reminders to all candidates on either the evening before or the morning of the debate with reminders and information on where to go.

By emphasizing the credibility of the debates and providing information about the structure of the debates, the additional invitation efforts served to reduce uncertainty about the returns to debate participation. As a result, while the intervention was administered to all candidates in treated districts, as highlighted by our pre-analysis plan, we expected it to be more likely to affect the participation decision of leading candidates for the reasons discussed above. Among these candidates, we expected that the intervention would particularly affect the participation of those who were most uncertain about the return to debate participation.

We also cross-randomized the extent of debate rebroadcasting, which ultimately had no effect because citizens were likely to hear their district debate even in districts without intensive rebroadcasting, likely reflecting the level of citizen interest in the debates. Because candidates were unaware of any differences in future rebroadcasting efforts, we present results where we pool over rebroadcasting intensity for clarity of exposition. According to our candidate survey, candidates believed that debates would be rebroadcast roughly two times, with no statistical difference by treatment assignment.

[12]Randomizing invitation effort at the candidate level would have raised serious concerns in terms of ethics and fairness to candidates.

[13]USAID was the donor funding the debate initiative. We expected their branding to be persuasive since our candidate survey indicates that more than 43% of the candidates report having worked for an international NGO.

[14]Candidates were informed that questions would be asked about relevant local policy issues but were not provided specific questions.

3.2 Debate participation

While the debates were well-attended overall, there was substantial variation across districts, ranging from 11% to 100%.[15] Overall 59% of candidates participated, which broke down to 48% of incumbents compared to 60% of challengers and other candidates. As shown in Table 3, the reasons cited for participation by our candidate survey respondents varied. Non-incumbent candidates cited their democratic duty more than any other reason, while for incumbents the most common reason was the opportunity to showcase their policy platforms to voters. All candidates viewed the debates as an opportunity to publicize their campaigns, but challengers put more emphasis than other candidates on the opportunity to demonstrate their competence. Only a small share of candidates pointed to radio broadcasting as a reason for their participation or admitted to attending in order to attack other candidates. These differences suggest a greater degree of political sophistication among incumbent candidates, who prioritized the instrumental value of the debates in publicizing their policy priorities.

3. Reasons Cited For Debate Participation

Candidate type	Duty (1)	Policies (2)	Competence (3)	Publicity (4)	Radio (5)	Attack (6)
Incumbent	0.40	0.80	0.27	0.40	0.07	0.07
Challenger	0.61	0.48	0.37	0.35	0.02	0.07
Other	0.54	0.52	0.25	0.43	0.01	0.09

The table displays mean values of binary column variables among incumbents, challengers and other candidates based on our surveys. The candidates were allowed to cite more than one reason for debate participation. 'Duty:' cited democratic duty. 'Policies:' cited opportunity to present policy platform. 'Competence:' cited opportunity to show off competence. 'Publicity:' cited opportunity for free campaign publicity. 'Radio:' cited the benefits of radio broadcasting reaching a large audience. 'Attack:' cited opportunity to attack other candidates.

Our candidate survey is also informative about the reasons why candidates were not present at their debate, although candidates predictably cite logistical issues rather than the electoral risks they faced in participation. Over 50% of non-participating candidates cited late notice, while 30% claimed that they did not receive an invitation. Nearly 20% mentioned road conditions to justify their absence.

Next, we analyze the debates themselves. Using transcript data we confirm that, on the whole, the unbiased rules of debate moderation were kept and candidates were given equal time to outline their policy priorities (see Table 4). Candidates varied in how they emphasized

[15]Since legislative campaigns are only loosely overseen by the relatively weak national parties, we found no evidence of parties coordinating the debate decisions of their candidates across districts or in response to the intervention.

their own qualifications during introductions, with incumbents focusing on their experience while challengers and other candidates highlighted their educational achievements. The most commonly cited policy priorities related to district primary schools, health facilities, and the quality and extent of roads. However, incumbents spoke at greater length about both the County Social Development Fund (CSDF) and the Legislative Support Project (LSP) funds, reflecting their first-hand experience with these funds. Finally, incumbents were much more likely to both be attacked by other candidates and attack others, as their attendance seemed to act as a focal point for other debate participants.

4. Transcript Descriptive Statistics

Candidate type	Intro words (1)	Education emphasis (2)	Experience emphasis (3)	CSDF words (4)	LSP words (5)	Attacked (6)	Attacker (7)
Incumbent	340.3	0.22	0.33	398.2	224.0	0.19	0.15
Challenger	352.0	0.30	0.26	284.7	218.0	0.04	0.04
Other	345.9	0.27	0.19	269.8	203.7	0.03	0.03

Table displays mean values of column variables among incumbents, challengers and other candidates based on our surveys. 'Intro words': number of words spoken in introduction. 'Education emphasis': candidate highlighted their education in introduction. 'Experience emphasis': candidate highlighted their experience in introduction. 'CSDF words': number of words spoken about ways to improve management of County Social Development Funds. 'LSP words': number of words spoken about priorities for spending Legislative Support Projects funds. 'Attacked': candidate was verbally attacked by another candidate. 'Attacker': candidate verbally attacked another candidate.

Qualitative evidence from focus groups suggests that citizens were affected by candidate participation in the debates and found the presentation of concrete policy platforms to be novel. As one participant said, "Before the debate, the word 'platform' was a strange word to me" (Vai Town, 26 September 2017). Many commented that the debates increased information available about candidates, noting that "in the past, there was no opportunity created for voters to engage candidates in understanding their platforms" (Foya, 20 September 2017). As a result, it is not surprising that citizens took note of participation decisions, highlighting that "we wanted to see all the six candidates at this debate but only two appeared, which is not good because we are not hearing from [the] other four candidates" (Massabolahun, 21 September 2017). Some even wanted debates to be mandatory: "There should be a law binding all candidates to attend the debate... You can't be somebody who wants to represent me if you don't turn up" (Vai Town, 26 September 2017).

Our qualitative evidence also suggests that the debates caused voters to change their assessment of the candidates. One participant stated, "The debate changed my attitude toward

candidates and helped me discover the hidden secret of some candidates" (Kolahun, 18 September 2017). Similarly, another participant mentioned, "For me, when I reached there, the first person I wanted to vote for ... well, my mind did not go on him. When I entered inside the debate and heard them speak my mind started going on another candidate" (Klay, 27 September 2017). Several focus groups pointed to the lack of specificity in candidates' promises and the mixed quality of policy platforms. As an example, one participant indicated that "some of the candidates were not detailed in their explanation on how they going to tackle these sectors" (Voinjama, 12 September). In particular, some respondents argued that the policy platforms of challengers were often weaker than those of the incumbents, as exemplified by one person's comment that "I did not hear anything new from candidates contesting against the incumbent because the incumbent was already doing most of these things" (Kolahun, 18 September 2017).

4 Data

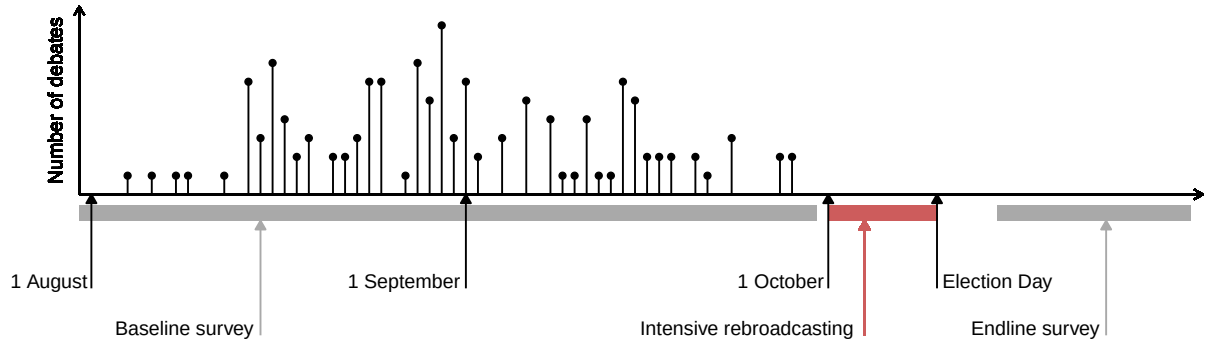
Our primary data source is a panel survey of 4,060 registered voters conducted in all 73 electoral districts in the country. In these interviews, enumerators used tablet computers while making phone calls to respondents sampled from the universe of active cell phone numbers on the country's largest mobile network. The distribution of observations per electoral district naturally reflects cell phone penetration and rurality. As the descriptive statistics shown in Table A2 indicate, the sample is older, more male, and better educated than the average Liberian.

In Figure 2 we show a timeline of the debates and data collection. Our survey began in early August, right before the first debates. Most data collection was completed by early September but concluding the baseline survey in several electoral districts took several more weeks.[16] The overlap of the baseline survey and the live debates is not a major concern. First, we control for any potential baseline debate exposure using the date on which respondents were interviewed. Second, for variables that were collected only for the endline survey, the timing of the baseline survey is irrelevant. Lastly, the intensive rebroadcasting of debates took place beginning October 1, by which time 88% of baseline data had been collected.

We use several other data sources. First, we use polling station-level election results to assess effects of the intervention on administrative electoral outcomes. Second, we conducted a survey of more than 600 of the candidates who ran in the election. We employ this data to

[16]Since the cellphone number sample was stratified at the county-level, sampling in some districts proved difficult, especially when the county contained both urban and rural districts since, in these cases, most calls went to those in the urban districts, and so achieving a sufficient sample in the rural districts took longer than anticipated.

Figure 2. Timeline of Debates Initiative



provide descriptive evidence on candidacy, validate important aspects of the intervention, and provide further evidence of the mechanism behind our results. Third, we use debate transcripts from each debate for descriptive evidence on the debates themselves.[17] Fourth, we conducted a survey of more than 50 radio stations to gather descriptive evidence about the stations and to validate their frequency of rebroadcasting the debates.

4.1 Outcome variables

To assess whether the invitation intervention, the debates themselves, and the debate rebroadcasting were properly implemented, we exploit two main pieces of data. For candidate debate participation, we use data from administrative debate reports as well as debate transcripts. For radio rebroadcasting, we use data from the rebroadcast schedules contracted with each of the radio stations and Naymote, a youth organization, which was hired to tune into each scheduled transmission to ensure debates were being played, unedited and on schedule. We complement this data with responses to our survey of radio stations to assess whether contracted and non-contracted stations also rebroadcast the debates or related content at other times.

To measure our key outcome variables, we rely on our voter survey and polling station-level data. For all outcome variables, we provide general descriptions in the relevant regression tables while details on their construction are in Appendix A.4. Whenever relevant, we aggregate related outcome variables using standardized z-scores as described by Kling et al. (2007).

To evaluate the ultimate electoral consequences of the intervention, we use the self-reported vote choices of respondents and validate these results using polling station-level electoral results.

[17] Internews partnered with the *Daily Observer* newspaper and hired trained journalists to transcribe each debate.

To measure debate exposure and information acquisition, we asked respondents questions about the debates, policy issues discussed within the debates, and about discussions they held with others about the debates. To assess individual beliefs about the policy priorities and competence of candidates in their districts, we ask respondents about both such perceptions and associated uncertainty, but only about three *predicted leading candidates*, as described in Section 2.1 and explained fully in Appendix A.1.[18] We ask individuals about their exposure to the campaign efforts of each of these predicted leading candidates as well as their beliefs about the candidates' competence and policy platforms. For all respondent-candidate dyads, we split the analysis into the incumbent and a pooling of the predicted challengers. We provide descriptive statistics for all the raw respondent level outcome variables in Table A3 and for all the raw respondent-candidate level outcomes in Table A4.

Because we had no control over the local issues that would be discussed in the debates, all questions about policy priorities were open-ended. These include both the questions about the policy priorities of the voters as well as their perceptions about the priorities of the *predicted leading candidates*. To analyze these priorities, open survey responses were coded by independent coders with no knowledge of treatment assignment.

4.2 Interaction and selection variables

As stated in our pre-analysis plan, we expect that voting outcomes will be affected by two key variables: candidate performance in the debate and the extent of alignment between citizens' and candidates' policy priorities. In our survey analysis, we asked citizens about who they thought won their district debate. This performance measure, however, is naturally correlated with treatment assignment and so in the main analysis we construct a *predicted* candidate-level measure of debate performance,[19] and use this to predict performance for the full sample of candidates irrespective of actual participation.

We measure the extent of preference alignment between respondents and candidates using data from our baseline survey in which we asked respondents to name their top three policy priorities in their district as well as to name what they believe are the top three policy priorities for each of the three predicted leading candidates. We aggregate this latter measure across

[18] This is both because asking about up to 28 candidates would have made the survey prohibitively time-consuming and because we had theoretical reasons to expect that the invitation intervention should affect the attendance decision of the most relevant candidates to a greater degree than it would affect other candidates.

[19] We do this by flexibly estimating a LASSO model of the debate performance outcomes of participating candidates on pre-determined covariates including their gender, incumbency status, party, baseline citizen assessments of their competence, policy priorities, and certainty regarding both competence and policy priorities.

respondents to the district level to create a measure of each candidate's policy priorities. We then calculate the share of a given respondent's top issues that are shared with each candidate to create a measure of preference alignment. We also create a version where we calculate the average of this variable at the district-level.

To assess differential candidate selection into debate participation, we also use such district-level measure of preference alignment, along with a secondary measure where we instead base candidates' priorities on their own survey responses. This alternative measure, as we discuss in Section 7.1, suffers from the post-election timing of the enumeration as well as only being observed for the 60% of candidates who responded to our survey. Lastly, we additionally consider citizens' certainty about candidate policy priorities, also aggregated to the district level.

5 Estimation

We randomly assigned all 73 districts into one of two treatment conditions: low invitation effort (the control group) or high invitation effort (the treatment group). To assign treatment conditions, we first pre-stratified based on which of the debate partners was running that district's debate. This is because the capacity of the debate organizers varied substantially in terms of their ability to attract candidates and organize the logistics of the debates. Second, we blocked on a set of pre-treatment covariates at the district-level to maximize power.[20] This strategy generated 19 blocks, with 3 or 4 districts per block. We then randomly assigned 38 districts to the treatment and 35 to control. Pre-treatment covariates at the district, individual, polling station, and candidate levels are well-balanced across treatment conditions. Full details are provided in Appendix A.2.

5.1 Estimating equations

Taking the case where the respondent-candidate is the unit of observation, we estimate:

$$y_{icdeb} = \beta T_{db} + \mathbf{X}_{icdb} + \mathbf{Z}_{db} + \theta_e + \eta_b + \epsilon_{icdeb}, \tag{1}$$

where y_{icdeb} is the outcome for respondent i regarding candidate c in district d interviewed by enumerator e in block b.[21] T_{db} is an indicator for districts assigned to treatment. η_b are

[20] We blocked on variables described in Panel A of Table A2.

[21] This estimation approach extends to cases where the respondent is the unit of observation, y_{ideb}, and where the

randomization block fixed effects and θ_e are survey enumerator fixed effects. Throughout, we include both district-level covariates $\mathbf{Z}_{db}$ and individual-level covariates $\mathbf{X}_{icdb}$ for which we provide descriptive statistics in Panels A and B of Table A2, respectively. While we have good balance on these covariates we include them to improve precision. We cluster standard errors at the district level, our level of treatment assignment.

At the individual level, we report pre-registered specifications that vary the weighting of observations to help account for variation in the number of survey respondents by electoral district. We report specifications where observations are unweighted; weighted by the inverse of the number of respondents in that district-wave (1/Obs); and weighted by the number of registered voters in that district divided by the number of respondents in that district-wave (Reg/Obs).The '1/Obs' weights imply that each district overall is equally weighted in the estimation, while the 'Reg/Obs' weights imply that districts are weighted in proportion to their share of registered voters. In the Appendix, we provide an additional set of results for all tables with survey-based outcomes where we instead weight observations to be representative of district-level demographics.

We consider an analogous specification to that of Equation (1) for electoral outcomes available at the polling station-level. In this case, instead of individual-level controls, we control for polling station-level variables listed in Panel C of Table A2. At the polling station-level, we report specifications that are unweighted, with observations weighted by the inverse of the number of polling stations in that district (1/PS), and by the number of registered voters at that polling station (Reg). The weighted specifications help account for variation in the number of polling stations by electoral district.

Whenever we have a panel for a given question where the outcome is continuous, we consider the continuous change in that variable between baseline and endline as an outcome Δy_{icdeb}. When the outcome is binary, we construct an indicator for whether the coded response changed between waves. The estimating equation remains the same aside from controlling for whether respondents were interviewed at baseline before or after the first broadcast of their district debate and its interaction with treatment assignment. Lastly, we also make use of specifications where we interact treatment assignment with candidate-level covariates X_{cdb}, which applies to the interactions discussed in Section 4.2.

Our coefficient of interest in Equation (1) is β, which recovers the reduced-form or intent-to-treat treatment effect of the invitation intervention. While the treatment effect of debate

candidate is the unit of observation, y_{cdb}.

participation, or the instrumental variable estimate that we would recover by instrumenting debate participation with assignment to treatment, is of great theoretical interest, some of our estimates indicate that the exclusion restriction condition that is required for the IV estimate to be unbiased is unlikely to hold.

5.2 Presentation of results and pre-analysis plan

In our main results, first, we estimate the effect of the invitation intervention on debate participation by candidate type. Second, we estimate effects on electoral outcomes by candidate type, and assess hetereogeneity by policy priority alignment and predicted debate performance. These analyses are each consistent with our pre-analysis plan, which predicted that treatment effects on increased debate participation would be concentrated among leading candidates and that the effects on electoral support would be concentrated on candidates with better policy priority alignment and predicted debate performance. We did not pre-register hypotheses that differentiated effects between incumbents and challengers, or about the main effect of treatment assignment on electoral outcomes by candidate type.

To parse the results on debate participation and electoral support by candidate type, we estimate secondary treatment effects on voter and candidate response. First, we assess differences in compliance with the intervention between incumbents and challengers. Second, we estimate treatment effects on citizen attention to the debates and information acquisition, which we hypothesized would be greater in treated districts. Third, we estimate effects on citizens' updating about the candidates, which we also hypothesized would be greater in treated districts. Fourth, we estimate effects on candidates' campaigning intensity. We hypothesized that overall on-the-ground campaigning would not increase in treated districts, due to the difficulty of spatially targeting campaigning responses after the broad dissemination of their policy promises. We document and justify divergences from our pre-analysis plan in Appendix A.3.

6 Main results

In this section we establish our two main results. First, we show that the intervention increased the likelihood that leading candidates, whether incumbents or challengers, attended their debates in treated districts. Second, we show that incumbents, especially those whose policy priorities aligned with those of their constituents and we predicted to perform well at their

debates, electorally benefited in treated districts at the expense of their challengers.

6.1 Effects on debate participation

First, we assess whether the intervention had its intended effect on debate participation.[22] Table 5 reports treatment effects on the attendance of candidates at the debates.[23] Column 1 in Panel A suggests that the invitation intervention led to a 7.7 percentage point (pp) (14% relative to the control mean) increase in the share of total candidates attending the debates in treated districts. In Panel B we show that incumbents were 21.2 pp (76%) more likely to attend in treated districts, and in Panel C we show that predicted challenger candidates were 21.2 pp (43%) more likely to attend. Lastly, Panel D reports no treatment effect on other candidates, reinforcing our expectation that the invitation intervention would mainly affect the debate participation decision of leading candidates.

6.2 Effects on electoral outcomes

Second, we establish that the intervention ultimately affected voting behavior.[24] Table 6 presents treatment effects on vote choice, using both our citizen survey (Panel 1) and polling station electoral returns (Panel 2). Panel 1 presents a set of specifications testing for whether there were treatment effects on vote choice, defined as whether the respondent reported voting for a given candidate at endline.[25] Columns 1-3 present the main effects of the treatment assignment. Columns 4-6 present specifications in which we interact treatment assignment with the standardized measure of predicted debate performance described in Section 4.2. In columns 7-9, we interact treatment assignment with the standardized measure of policy preference alignment between the citizen and the candidate described in the same section.

In Panel 1.A., focusing on the incumbent, we find strong positive treatment effects on vote choice. Incumbents experienced a 4.5 pp increase in respondent vote choice in treated districts.

[22] Appendix Table A20 shows similar results when we instead aggregate candidate participation to the district-level, but we focus on the respondent-level results for consistency across first-stage and reduced-form results. Table A21 further shows that whether the eventual election winner attended and the share of actual challengers who attended was substantially affected by treatment assignment, as expected.

[23] In Appendix Table A22 we show that there is no difference in how often the debates were rebroadcast. This is measured either based on the radio monitors who tuned into each contracted rebroadcast, or based on our radio station survey that included radio stations not contracted to rebroadcast but recorded as present in a district debate.

[24] Appendix Table A23 reports that citizens in treated districts show no change in the reasons they cited for their vote choice in the baseline versus endline surveys, which thus cannot account for our results on voting.

[25] We pre-registered using an outcome variable defined as vote switching towards a given candidate. However, since few respondents indicated a concrete vote choice at baseline, using either measure produces qualitatively identical results. Focusing on endline responses allows us to compare voter responses to polling station outcomes.

5. Candidate Debate Participation

	(1)	(2)	(3)
A. Share of candidates			
Invite	0.077**	0.065**	0.092***
	(0.034)	(0.030)	(0.033)
Control Mean	0.542	0.573	0.557
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
B. Incumbent			
Invite	0.212**	0.177**	0.234***
	(0.083)	(0.073)	(0.083)
Control Mean	0.280	0.372	0.299
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
C. Share of challengers			
Invite	0.212***	0.144**	0.220***
	(0.074)	(0.063)	(0.067)
Control Mean	0.492	0.554	0.528
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
D. Share of other candidates			
Invite	0.003	0.008	0.009
	(0.030)	(0.028)	(0.029)
Control Mean	0.562	0.583	0.575
Observations	3991	3991	3991
Weight	No	1/Obs	Reg/Obs

Outcome variables are the share of the respective set of candidates (all, incumbent, predicted challenger, other candidate) who attended a debate out of all candidates in that district. Panels A-C have 4060 observations; Panel D has fewer due to only three candidates running in two districts (and hence no 'other candidates' defined).

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6. Voting Outcomes

	Main effect			Interaction term: Std. Performance			Interaction term: Std. Policy alignment		
1. Respondent-level	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Incumbent									
Invite	0.045**	0.051**	0.035*	0.049**	0.054**	0.039*	0.041**	0.045**	0.032
	(0.020)	(0.023)	(0.020)	(0.020)	(0.023)	(0.020)	(0.019)	(0.023)	(0.019)
Invite × Std. performance				0.067*	0.059	0.039			
				(0.038)	(0.036)	(0.031)			
Invite × Std. policy alignment							0.042***	0.051***	0.052***
							(0.016)	(0.017)	(0.018)
Control Mean	0.278	0.282	0.288	0.278	0.282	0.288	0.278	0.282	0.288
Observations	3496	3496	3496	3496	3496	3496	3496	3496	3496
B. Challengers									
Invite	-0.048***	-0.039***	-0.036**	-0.051***	-0.039**	-0.039**	-0.048***	-0.039***	-0.036**
	(0.014)	(0.013)	(0.014)	(0.017)	(0.015)	(0.016)	(0.014)	(0.013)	(0.014)
Invite × Std. performance				-0.015	-0.002	-0.026			
				(0.046)	(0.039)	(0.043)			
Invite × Std. policy alignment							-0.006	-0.004	-0.003
							(0.007)	(0.008)	(0.008)
Control Mean	0.156	0.156	0.156	0.156	0.156	0.156	0.156	0.156	0.156
Observations	8684	8684	8684	8684	8684	8684	8684	8684	8684
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs
2. Polling station-level	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Incumbent									
Invite	0.042*	0.037	0.041*	0.044*	0.036*	0.044*	0.044*	0.041*	0.043*
	(0.024)	(0.023)	(0.024)	(0.023)	(0.021)	(0.023)	(0.023)	(0.022)	(0.023)
Invite × Std. performance				0.082***	0.099***	0.081***			
				(0.027)	(0.030)	(0.026)			
Invite × Std. policy alignment							-0.016	-0.017	-0.017
							(0.028)	(0.029)	(0.029)
Control Mean	0.246	0.250	0.250	0.246	0.250	0.250	0.246	0.250	0.250
Observations	4618	4618	4618	4618	4618	4618	4618	4618	4618
B. Challengers									
Invite	-0.029**	-0.031**	-0.028**	-0.033**	-0.036***	-0.031**	-0.029**	-0.029**	-0.029**
	(0.012)	(0.012)	(0.012)	(0.013)	(0.012)	(0.013)	(0.012)	(0.013)	(0.012)
Invite × Std. performance				-0.034	-0.031	-0.037			
				(0.027)	(0.029)	(0.026)			
Invite × Std. policy alignment							-0.000	-0.009	0.005
							(0.019)	(0.017)	(0.019)
Control Mean	0.113	0.112	0.112	0.113	0.112	0.112	0.113	0.112	0.112
Observations	11385	11385	11385	11385	11385	11385	11385	11385	11385
Weight	No	1/PS	Reg	No	1/PS	Reg	No	1/PS	Reg

Panels 1.A. and 1.B: Outcome variable is an indicator for whether respondent expressed voting for either the incumbent (Panel A) or a predicted challenger (Panel B) at endline. Columns 1-3 report the main effects, Columns 4-6 include interactions of treatment assignment with standardized candidate-level measures of predicted debate performance, and Columns 7-9 include interactions with standardized respondent-candidate-level measures of policy preference alignment measured at baseline. Panels 2.A. and 2.B: Outcome variables are vote shares over the number of registered voters for the incumbent (Panel A) or predicted challengers (Panel B) using polling station-level electoral data. We use the district-level analogues of the interaction terms used in Panel 1. Section 4.2 explains these interaction terms further.

Specifications using citizen survey outcomes are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Specifications using polling station-level data are estimated using OLS and include block FE, district-level and polling station-level controls. For weighted specifications, 'PS' is the number of polling stations in that district and 'Reg' is the number of registered voters at that polling station. Standard errors clustered at the district-level in parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Moreover, the significantly positive interactions with both debate performance and policy priority alignment indicate that this effect is greater among incumbents who were predicted to perform well in the debates and whose policy priorities align with those of their electorates. In contrast, focusing on predicted challengers in Panel 1.B., there are broadly negative main effects and little evidence of significant interactions. Challengers experienced a significant 4.8 pp drop among respondents' vote choice in treated districts. Importantly, this negative effect on challengers' voting outcomes is not entirely mechanically implied by the positive effect on incumbent voting outcomes, since our categorization of *predicted challengers* excludes the majority of more marginal candidates.

In Panel 2 of Table 6, we use polling station-level data to validate our survey results. These results reinforce our findings. In Panel 2.A., we find that incumbent vote share in treated districts was 4.2 pp higher in treated districts than in control districts, albeit somewhat more noisily estimated. Incumbent vote share is larger in districts where their predicted debate performance was higher. However, we find no significant interaction with our district measure of policy priority overlap here, which we attribute to the fact that our sample of district respondents is not representative and thus might provide a noisy measure of the the alignment between candidates' policy priorities and all voters in the district. In Panel 2.B., we continue to find that challenger vote share in treated districts is 2.9 pp smaller than in control districts, and there is no evidence of interactions with predicted debate performance or overlap in policy priorities between challengers and citizens in their districts. The polling station results broadly support the survey results and suggest positive electoral consequences for incumbents in treated districts but negative consequences for their challengers.

7 Explaining the results

Our main results suggest that electoral gains accrued to incumbent candidates in treated districts, particularly when they were predicted to perform well in their debates and their policy priorities aligned with those of their constituents. These results mirror actual election outcomes: 50% of incumbents in treated districts won re-election compared to 43% in control districts.[26] In a context where approval of incumbent performance is generally low, and given the results of prior experimental interventions, these results might seem surprising. In this section, we establish three sets of secondary results that together explain these electoral effects: variation in

[26]This difference is not surprising given that more than 35% of races are decided by winning margins of less than five percentage points, which approximates the treatment effects on voting outcomes shown in Table 6.

compliance with the intervention by different type of candidates; effects on the attention paid by voters to the debates and belief updating about candidates; and how candidates' campaigning responded. We discuss each in turn.

7.1 Differential selection into debate participation

Understanding the effects on electoral outcomes demands an analysis of compliance with the invitation treatment and hence candidate selection into the debates themselves. Few incumbents attended in control districts—just 35%—whereas the additional invitation effort increased this number to around 50% in treated districts. We show that incumbents, both always takers and compliers, self-selected into debate participation, in terms of the degree of their policy priority alignment with citizens in their districts, more positively than challenger candidates did.

We consider standardized candidate-level measures of (1) policy alignment with citizens in the district; (2) citizens' certainty about candidate policy priorities. We follow Abadie (2003) and Angrist and Fernandez-Val (2013) to compute the kappa-weighted means of these variables for the various compliance groups. Table 7 presents these estimates when employing the baseline citizen survey to construct the variables in Panel A, and when employing the candidate survey to construct the policy alignment variable in Panel B.[27] We focus on the estimates based on the baseline citizen survey. This is because our candidate survey is missing a substantial share of responses, differentially across candidate type, and was conducted after the election due to the logistical constraints that entailed surveying candidates while they were campaigning.[28] Nonetheless, estimates using both data sources provide qualitatively similar results.

Considering policy alignment, *on average* incumbents are very similarly aligned with citizens compared to challengers. However, we only find a positive pattern of self-selection into debate participation among incumbents. Always-taker incumbents were very well aligned with the policy priorities of citizens in their districts, while complier incumbents were aligned better than never-taker incumbents. Among challengers, we find a much less clear pattern. While always-takers do not seem to self-select relative to never-takers based on their policy alignment

[27] As explained in Section 4.2, the policy alignment measure based on the citizen survey measures the overlap between the top three priorities by the candidates, as perceived by citizens as baseline, and the top priorities reported by citizens in their district. The citizens' certainty measure is the average reported by citizens in their district. In turn, the policy alignment measure based on the candidate survey measures the overlap between the top three priorities reported by the candidates themselves and the top priorities report by citizens in their district. There is no analog of the citizens' certainty measure when using the candidate survey.

[28] Overall response rates are 47% among incumbents, with compliers responding 50% of the time. Overall response rates are 65% among challengers, with compliers responding only 47% of the time. These differential response rates by compliance status might bias the estimates based on the candidate survey data.

7. Characterizing Compliers

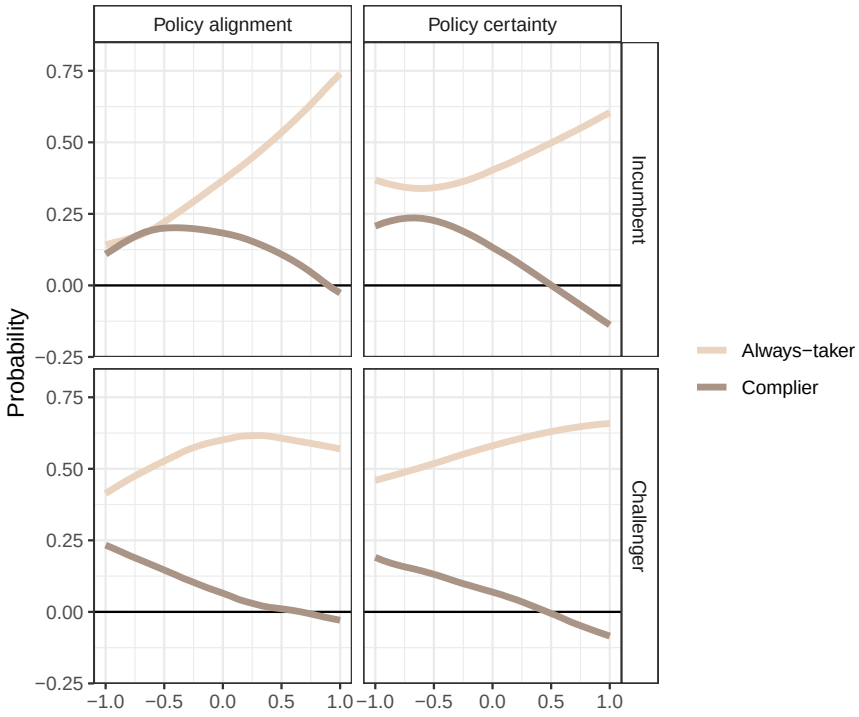
	A. Citizen survey				B. Candidate survey			
	All	C	AT	NT	All	C	AT	NT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Policy alignment								
Incumbent	-0.08	-0.31	0.57	-0.58	0.07	1.06	0.86	-0.47
Challengers	0.03	-1.31	0.19	0.16	-0.02	-1.72	0.08	0.10
Policy certainty								
Incumbent	0.36	-0.76	0.52	0.43				
Challengers	-0.14	-1.23	-0.04	-0.06				

Each variable is standardized. Panel A uses citizen survey to construct variables; Panel B uses candidate survey to construct policy alignment variable. Mean of each variable presented for all candidates (All); compliers (C); always-takers (AT); never-takers (NT). Calculations follow Abadie (2003) and Angrist and Fernandez-Val (2013).

with citizens, compliers *negatively* do. Considering certainty about policy priorities, on average citizens were much more certain about incumbents' priorities than challengers'. Consistent with the idea that the intervention allayed concerns about the risk of debate attendance, there was less certainty about the policy priorities of complier candidates than either always-taker or never-taker candidates.

In Figure 3 we nonparametrically estimate the probability of a given candidate being an always-taker or complier across different values of these variables. The plots in the top panel corroborate a strongly positive self-selection among always-taker incumbents based on their policy alignment with constituents. Among candidates in districts assigned to the control, incumbents with policy alignment 1 sd above the mean participated at rates of around 75%. Those with policy alignment 1 sd below mean participated less than 20% of the time. Consistent with Table 7, the plots suggest that the intervention induced the participation of incumbents at intermediate levels of policy alignment. The plots in the bottom panel, in turn, confirm a substantially weaker self-selection of always-taker challengers on candidate policy alignment. With respect to complying challengers, in contrast with complying incumbents, we see a negative self-selection on policy alignment. However, as with complying incumbents, we observe that there is substantially less certainty about their policy priorities.

Figure 3. Complier Status by Baseline Candidate Characteristics



Note: Plot presents fitted values from a non-parametric regression of compliance status across standardized values of baseline characteristics for incumbents (top row) and challengers (bottom row). At a given value of each baseline characteristic, the fitted value indicates the probability of a candidate type being an always-taker or a complier. Optimal bandwidths computed using Calonico et al. (2018).

7.2 Voters paid more attention to the initiative

Next, we assess how voter exposure and information acquisition was affected by treatment assignment. In Panel A of Table 8, we use a standardized index of our measures of direct exposure to the debates, including whether respondents heard the debate and how often they heard them. The results indicate that citizens in treated districts had exposure 0.304 standard deviations (sd) higher than those in control districts.[29]

In Panel B, we use an index reflecting factual knowledge about the debates themselves. The results suggest that citizens in treated districts had factual knowledge of the debates 0.125 sd higher than citizens in control districts. Given that our endline survey began around a month after the election, the persistence of this effect suggests meaningful differences in debate exposure. In Panel C we use an index reflecting factual knowledge about a national policy issue, the management of County Social Development Funds, which was asked about in every debate. We find treatment effects of 0.156 sd on correctly learning about management issues of these poorly-understood funds. In Panel D we use an index reflecting broader political information acquisition relating to discussion about the debates and listening to the radio. The estimates suggest substantial effects on information acquisition corresponding to a 0.251 sd treatment effect.

These results provide strong support for the idea that the invitation intervention meaningfully affected citizen exposure to political information through the debates.[30] They suggest that the debates were relatively effective at conveying policy information, particularly when leading candidates participated, which voters in turn discussed with each other. We provide results disaggregating the components of these indices in Table A18. We additionally show, in Appendix A.5, that treatment assignment additionally affected levels of broader political engagement by citizens and the extent to which they report coordinating their vote choices.

Next, we show that the intervention, through increasing both candidate participation and thereby citizen interest and engagement with the debates, led voters to update about the competence and policy priorities of predicted leading candidates. We first assess treatment effects on the standardized change in citizens' reported certainty about the competence (columns 1-3) and priority issues (columns 4-6) of incumbents, shown in Panel 1.A. of Table 9. The results

[29] Importantly for our ability to pool districts assigned to different rebroadcasting intensity, as we show in Appendix A.3, there were no differential treatment effects on citizen exposure to the debates when pooling districts that were cross-randomized to different rebroadcasting intensities or when restricting to districts assigned to high invitation intensity.

[30] This is also consistent with the theory that more deliberation between relevant candidates will increase voters' attention (Wantchekon et al., 2018).

8. Debate Exposure and Information Acquisition

	(1)	(2)	(3)
A. Debate listening index			
Invite	0.304***	0.341***	0.424***
	(0.102)	(0.105)	(0.107)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
B. Debate knowledge index			
Invite	0.125**	0.126**	0.165***
	(0.063)	(0.058)	(0.059)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
C. Policy knowledge index			
Invite	0.156*	0.230*	0.189*
	(0.089)	(0.121)	(0.100)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
D. Political information acquisition			
Invite	0.251***	0.313***	0.300***
	(0.078)	(0.090)	(0.091)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs

Outcome variables: Panel A: a standardized index of (1) indicator for whether the respondent had not heard their district debate at baseline but had at endline (2) indicator for whether the respondent had heard the debate at endline (3) the number of times the respondent had heard the debate at endline. Panel B: a standardized index of (1) indicator for whether the respondent's stated debate winner actually attended the debate (2) share of candidates respondent claims participated (3) share of predicted leading candidates respondent claims participated. Panel C: a standardized index of the change in how many factual questions about CSDF management respondents answered correctly between baseline and endline. Panel D: a standardized index of (1) change in how much respondents listened to the radio (2) change in how much they discussed politics with their friends (3) how much they accessed other sources of political information.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

indicate that citizens in treated districts became significantly more certain about incumbent competence and priority issues, with standardized treatment effect sizes of 0.178 and 0.169 sd, respectively. In Panel 1.B. there is no evidence that citizens in treated districts became more certain about the competence of challengers, and some evidence that those citizens became more certain about challengers' priority issues.

We then assess treatment effects on the standardized change in citizens' beliefs about the competence (columns 1-3) and priority issues (columns 4-6) of predicted leading candidates, shown in Panel 2. These treatment effects suggest positive updating regarding incumbents and negative updating for their challengers, but the estimates are imprecise. Panel 2.A. reports suggestive sizable, but statistically insignificant at conventional levels, treatment effects that citizens feel more positive about incumbent competence and that they have learned more about their policy priorities. Specifically, for both outcomes, the treatment effect is around 0.1 sd. In contrast, Panel 2.B. indicates that citizens updated negatively about the competence of their challengers and they did not learn about the policy priorities of those candidates. The treatment effect on challenger competence is 0.1 sd and marginally significant ($p < 0.1$).

7.3 Campaigning response by candidates

Finally, we consider treatment effects on candidates' campaign efforts. In Table 10, we report results on standardized indices of survey responses regarding "on-the-ground" campaigning by candidates in respondents' towns (Columns 1-3) and "radio" campaigning (4-6). The on-the-ground campaigning index incorporates candidates' visits, distribution of leaflets, and vote buying in respondents' towns. The radio campaigning index captures candidate presence on the radio.

In Panel A, there is a significant increase in incumbent exposure on the radio in treated districts, but no significant treatment effect on on-the-ground campaigning by incumbents. Specifically, there is a positive treatment effect of 0.082 sd in incumbent radio campaigning. In contrast, Panel B reports evidence of negative treatment effects on on-the-ground campaigning by challengers, but no treatment effect on challenger radio exposure. Challengers reduced their on-the-ground campaigning by 0.060 sd in treated districts. We provide results disaggregating the on-the-ground campaigning index in Table A19.

9. Updating About Candidates

	Certainty about competence			Certainty about policy		
1. Uncertainty	(1)	(2)	(3)	(4)	(5)	(6)
A. Incumbent						
Invite	0.178*	0.186*	0.179**	0.169**	0.195**	0.192**
	(0.105)	(0.107)	(0.084)	(0.074)	(0.080)	(0.074)
Observations	3496	3496	3496	3496	3496	3496
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs
B. Challengers						
Invite	0.037	0.046	0.025	0.139**	0.118	0.098
	(0.066)	(0.074)	(0.070)	(0.061)	(0.073)	(0.067)
Observations	8684	8684	8684	8684	8684	8684
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

	Beliefs about competence			Learning about policy		
2. Levels	(1)	(2)	(3)	(4)	(5)	(6)
A. Incumbent						
Invite	0.098	0.066	0.093	0.089	0.125	0.091
	(0.075)	(0.089)	(0.082)	(0.065)	(0.089)	(0.073)
Observations	3496	3496	3496	3496	3496	3496
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs
B. Challengers						
Invite	-0.078	-0.147*	-0.096	0.038	0.027	0.063
	(0.075)	(0.088)	(0.077)	(0.065)	(0.086)	(0.080)
Observations	8684	8684	8684	8684	8684	8684
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

Panels 1.A and 1.B: the outcome variable in columns 1-3 is the standardized change in certainty respondents express about candidate competence between baseline and endline, and in columns 4-6 is the standardized change in certainty respondents express about candidate priority issues between baseline and endline. Panels 2.A and 2.B: the outcome variable in columns 1-3 is the standardized change in respondent perception about candidate competence between baseline and endline, and in columns 4-6 is the standardized learning that respondents reflect about candidate priority issues between baseline and endline.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

10. Candidate Campaigning

	Ground			Radio		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Incumbent						
Invite	-0.055	-0.075	-0.052	0.082**	0.087**	0.092**
	(0.043)	(0.050)	(0.051)	(0.037)	(0.040)	(0.042)
Observations	3492	3492	3492	3496	3496	3496
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs
B. Challengers						
Invite	-0.060*	-0.073**	-0.076**	-0.025	-0.004	-0.018
	(0.031)	(0.029)	(0.032)	(0.028)	(0.027)	(0.028)
Observations	8676	8676	8676	8684	8684	8684
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

Outcome variable in columns 1-3 is a standardized index of how often candidates (1) visited (2) distributed leaflets (3) bought votes in respondents' communities during campaigning. Outcome variable in columns 4-6 is a standardized measure of how often respondents heard candidates on the radio in the two weeks before the election.
All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

8 Discussion

Overall, the evidence indicates that incumbent candidates were much better at self-selecting into debate participation than challengers; they seemed better able to recognize when their policy priorities aligned with those of their constituents and when constituents were uncertain of this alignment. This effect is clear for both always-takers and compliers, which suggests that the treatment effects of the invitation intervention are likely representative of those of the overall debate initiative. Given the specific focus of the debates on policy issues, the results are overall consistent with incumbents being better equipped to correctly assess the returns to debate participation.[31] Challenger candidates, who frequently had less political sophistication and experience, appear to have been less successful at calculating the returns to participation.

When leading candidates were induced to participate, citizens paid more attention to the content of the debates and became more certain about the candidates. This increase in certainty was mainly for the incumbent, who spoke significantly more on the policy issue questions for

[31]It is worth considering whether this apparent greater sophistication is a product of prior selection or experience accrued over incumbency. While underpowered to conduct a regression discontinuity design to account for selection into incumbency, this exercise provides suggestive evidence that it is experience, rather than selection into incumbency, that accounts for our observed results.

which they possessed more experience. Similarly, the suggestive positive treatment effects on citizens' perceptions about incumbent competence but negative perceptions about challenger competence indicate that incumbents were likely to outperform challengers when they were induced to participate. These results align with both the differential patterns of selection into debate participation we observe, as well as the the focus group evidence presented in Section 3.2. The effects on voter response support the idea that incumbents but not their challengers benefited both in terms of increased voter information about their policy priorities and somewhat in terms of perceptions of their competence.

Finally, the effects on campaigning suggest that the presence of relatively sophisticated incumbents in the debates deterred challengers, who spent less time campaigning in person around their district in treated districts than did challengers in control districts. Moreover, incumbent debate performance led incumbents to increase their use of radio campaigning to complement the rebroadcasting of the debates. Suggestively, in our candidate survey, 77% of incumbents believed that the debates led radio stations to issue invitations for interviews more frequently, while only 63% of challengers believed they received more invitations. In addition, 68% of incumbents report that the debates led them to change their campaign strategy whereas only 52% of challengers reported a change. Restricting the data to candidates who said the debates affected their mode of campaigning, 60% of incumbents said they affected their radio appearances while only 40% said they affected their on-the-ground campaigning. By contrast, only 27% of challengers said the debates affected their radio appearances compared to 73% who said they affected their on-the-ground campaigning.

9 Conclusion

We examine the electoral consequences of candidates' decisions to supply programmatic policy information in a weakly-institutionalized, clientelistic setting. To that end, we randomly assigned an intervention to increase the participation of candidates in a nationwide debate initiative, which was designed to elicit and disseminate programmatic promises from candidates running for a seat in Liberia's House of Representatives in 2017. Policy promises were elicited from candidates who participated in the debates and disseminated via radio broadcasting.

We find that the intervention had uneven electoral consequences: incumbents who were induced to attend their debates by the intervention benefited at the ballot box at the expense of their challengers. These incumbents self-selected into the debates when their policy priorities

were aligned with their constituents, and thus they dominated their debates by proposing policy platforms closer to those preferred by their electorates than platforms espoused by their challengers. Our results suggest that both the debate content and the mode of dissemination mattered for this outcome. When relevant candidates attended their debates voters paid substantially more attention to the debate information and subsequently engaged in more political information acquisition. Incumbents induced to participate in the debates increased their radio exposure, aided by increased demand from radio stations, while their challengers reduced their on-the ground campaigning.

Our results point to the challenges of transitioning away from the low-accountability equilibrium characterizing many clientelistic democracies. By inducing a subset of self-selected candidates to compete on a more programmatic basis, our intervention had uneven electoral consequences that favored incumbents. In this context, there may exist substantial returns to incumbency, as challengers are usually less experienced and poorly equipped for programmatic competition. Incumbents, meanwhile, may possess informational advantages enabling them to better evaluate the return to participation in such initiatives. If such democratic initiatives are to enhance competition and contribute to shifting from a clientelistic towards a more programmatic political equilibrium, then they must at least tackle two constraints. First, given the implausibility of its enforcement, candidate incentives need to be aligned towards participation. Second, less experienced candidates should receive additional training to face a more leveled playing field. How to best achieve this remains a question for future research.

References

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics*, 113(2):231–263.

Afrobarometer (2015). Afrobarometer Data, Liberia, Round 6, 2014.

Al-Ubaydli, O., List, J. A., LoRe, D., and Suskind, D. (2017). Scaling for economists: Lessons from the non-adherence problem in the medical literature. *Journal of Economic Perspectives*, 31(4):125–144.

Angrist, J. D. and Fernandez-Val, I. (2013). Extrapolate-ing: External validity and overidentification in the late framework. In *Advances in Economics and Econometrics: Volume 3, Econometrics: Tenth World Congress*, volume 51, page 401. Cambridge University Press.

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M., and Walton, M. (2017). From proof of concept to scalable policies: Challenges and solutions, with an application. *Journal of Economic Perspectives*, 31(4):73–102.

Banerjee, A. V., Kumar, S., Pande, R., and Su, F. (2011). Do informed voters make better choices? experimental evidence from urban india. Working paper.

Barro, R. J. (1973). The control of politicians: An economic model. *Public Choice*, 14(1):19–42.

Bhandari, A., Larreguy, H., and Marshall, J. (2019). Able and mostly willing: An empirical anatomy of information's effect on voter-driven accountability in senegal. Working paper.

Bidwell, K., Casey, K., and Glennerster, R. (2019). Debates: Voting and expenditure responses to political communication. *Journal of Political Economy*, forthcoming.

Blattman, C., Larreguy, H., Marx, B., and Reid, O. (2019). Eat widely, vote wisely? Lessons from a campaign against vote buying in Uganda.

Bobonis, G., Gertler, P., Gonzalez-Navarro, M., and Nichter, S. (2019). Vulnerability and clientelism.

Bowles, J., Larreguy, H., and Liu, S. (2019). How weakly institutionalized parties monitor brokers in developing democracies: Evidence from post-conflict liberia. *American Journal of Political Science*, forthcoming.

Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2018). On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association*, pages 1–13.

Chong, A., De La O, A., Karlan, D., and Wantchekon, L. (2015). Does corruption information inspire the fight or quash the hope? a field experiment in Mexico on voter turnout, choice and party identification. *Journal of Politics*, 77(1):55–71.

Cruz, C., Keefer, P., Labonne, J., and Trebbi, F. (2019). Making policies matter: Voter responses to campaign promises. Working paper.

Cruz, C., Labonne, J., and Querubín, P. (2018). Politician family networks and electoral outcomes: Evidence from the philippines. *American Economic Review*, 107(10):3006–3037.

de Figueiredo, M. F., Hidalgo, F. D., and Kasahara, Y. (2013). When do voters punish corrupt politicians? experimental evidence from brazil. Working paper.

Duarte, R., Finan, F., Larreguy, H., and Schechter, L. (2019). Networks, information and vote buying. *Working paper*.

Dunning, T., Grossman, G., Humphreys, M., Hyde, S., McIntosh, C., and Nellis, G. (2019a). *Metaketa I: The Limits of Electoral Accountability*. Cambridge University Press.

Dunning, T., Grossman, G., Humphreys, M., Hyde, S. D., McIntosh, C., Nellis, G., Adida, C. L., Arias, E., Bicalho, C., Boas, T. C., Buntaine, M. T., Chauchard, S., Chowdhury, A., Gottlieb, J., Hidalgo, F. D., Holmlund, M., Jablonski, R., Kramon, E., Larreguy, H., Lierl, M., Marshall, J., McClendon, G., Melo, M. A., Nielson, D. L., Pickering, P. M., Platas, M. R., Querubín, P., Raffler, P., and Sircar, N. (2019b). Voter information campaigns and political accountability: Cumulative findings from a preregistered meta-analysis of coordinated trials. *Science Advances*, 5(7).

Enríquez, J. R., Larreguy, H., Marshall, J., and Simpser, A. (2020). Online political information: Facebook ads, electoral saturation, and electoral accountability in mexico. Working paper.

Fearon, J. D. (1999). Electoral accountability and the control of politicians: Selecting good types versus sanctioning poor performance. In Przeworski, A., Stokes, S., and Manin, B., editors, *Democracy, Accountability, and Representation*. Cambridge University Press.

Ferejohn, J. (1986). Incumbent performance and electoral control. *Public Choice*, 50(1):5–25.

Ferraz, C. and Finan, F. (2008). Exposing corrupt politicians: The effects of brazil's publicly released audits on electoral outcomes. *Quarterly Journal of Economics*, 123(2):703–745.

Finan, F. and Schechter, L. (2012). Vote-Buying and Reciprocity. *Econometrica*, 80(2):863–881.

Fujiwara, T. and Wantchekon, L. (2013). Can informed public deliberation pvercome clientelism? experimental evidence from benin. *American Economic Journal: Applied Economics*, 5(4):241–255.

Glencorse, B. and Yealue, L. (2017). Liberia's democracy is failing its people. *Washington Post*.

Global Witness (2016). The deceivers.

Hicken, A., Leider, S., Ravanilla, N., and Yang, D. (2018). Temptation in Vote-Selling: Evidence from a Field Experiment in the Philippines. *Journal of Development Economics*, 131(March):1–14.

Holmstrom, B. (1999). Managerial incentive problems: A dynamic perspective. *Review of Economic Studies*, 66(1):169–182.

IREDD (2016). Salaries and allowances of lawmakers in the national budget.

Kamara, K. A. (2017). Media strengthening, public participation and democracy. In Garnett, T., editor, *Liberia Development Conference Anthology*, pages 176–191. USAID.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

Larreguy, H. A., Marshall, J., and Snyder, James M., J. (2019). Publicizing malfeasance: When the local media structure facilitates electoral accountability in Mexico. *Economic Journal*, forthcoming.

Muralidharan, K. and Niehaus, P. (2017). Experimentation at scale. *Journal of Economic Perspectives*, 31(4):103–124.

National Democratic Institute (2014). Organizing and producing candidate debates: An international guide.

Platas, M. and Raffler, P. (2019). Closing the gap: Information and mass support in a dominant party regime. Working paper.

Spatz, B. J. and Thaler, K. M. (2018). Has liberia turned a corner? *Journal of Democracy*, 29(3):156–170.

USAID (2015). Liberia Electoral Access and Participation (LEAP) Survey. *United States Agency for International Development*.

USAID (2018). Liberia Electoral Access and Participation (LEAP) Survey. *United States Agency for International Development*.

Vasudevan, S. (2019). Diminishing the effectiveness of vote buying: Experimental evidence from a persuasive radio campaign in india. *Working paper*.

Vicente, P. C. (2014). Is Vote Buying Effective? Evidence from a Field Experiment in West Africa. *The Economic Journal*, 124(574):F356–F387.

Wantchekon, L., Lopez-Moctezuma, G., Fujiwara, T., Lero, C. P., and Rubenson, D. (2018). Policy Deliberation and Voter Persuasion: Evidence from an Election in the Philippines. *American Journal of Political Science*, 131(March):1–14.

A Appendix for Online Publication

A.1 Classifying candidates as leading candidates

For each candidate we constructed an indicator variable for whether the candidate was a *predicted leading candidate.* We constructed this indicator as follows, in a sequential fashion until there were three per district: (1) if a candidate was the incumbent; (2) if the candidate ran in the 2011 election and placed 2nd or 3rd; and (3) if the candidate was from a top party. We defined top parties as, sequentially, the incumbent Unity Party (UP), Coalition for Democratic Change (CDC), Liberty Party (LP), the Alternative National Congress (ANC) and the All Liberia Party (ALP). This process resulted in three selected candidates in all districts. These *predicted leading candidates* are then split into two groups: whether the candidate is the *incumbent* or whether they are a *predicted challenger*, i.e., a non-incumbent predicted leading candidate. The incumbent ran in 64/73 (88%) of races, and so in the remaining 9 districts all three of these candidates are coded as challengers. One additional incumbent ran in a new district and is consequently coded as a challenger. Validating our indicator for top candidates with actual electoral results, we find that in 50% of cases our predicted leading candidates came in the top three in their district, and in 71% of cases came in the top five. Given our aim to identify a set of relevant candidates who had plausible chances at electoral success and voters would be interested in, we consider the exercise to be successful.

We show in Table A21 that, using the 'actual' leading candidates who placed in the top three in the election—whether *actual leading candidates, winner* or *actual challengers*—generates a similarly strong first stage on debate participation. Using this alternative categorization generates a set of qualitatively similar results, albeit with a more restricted sample of only those candidates who were both *predicted* and *actual* leading candidates in the citizen-candidate level analysis. However, given the effects we find on voting outcomes, we consider it likely that the definition of *actual leading candidates* is endogenous to our intervention. These additional results are available on request.

A.2 Balance

We report balance on pre-treatment covariates at the district, individual, polling station and candidate levels. Balance is assessed by estimating Equation (1) for each covariate as an outcome, but omitting the individual-level $\mathbf{X}_i$ and district-level $\mathbf{Z}_d$ as controls. Across the different specifications, we present the coefficient on the treatment indicator *Invite* to test for evidence of imbalance between treatment groups. For district-level specifications, we report

specifications where districts are unweighted and where we weight by the number of registered voters, and use heteroskedasticity-robust standard errors. For individual-level specifications, we restrict to the the endline survey sample and consider two types of outcomes. First, we assign district-level outcomes to individuals in this sample. Second, we use individual-level covariates collected in the survey itself. We use weights as described in Section 5.1 and cluster at the district level. For the polling station-level specifications, first we assign district-level outcomes to each polling station in that district and second we use polling station-level variables using the fact that 90% of polling places in 2017 were also used in the 2011 election. For all new polling places we assign district-level averages. We present an unweighted specification, one where we weight by the inverse of the number of polling stations in that district and one where we weight by the number of registered voters in that polling station, and again cluster at the district level. Lastly, for the candidate-level specification we present an unweighted specification assessing balance on characteristics drawn from our candidate survey, as well as a weighted specification where we weight by the inverse of the number of responding candidate types in a given district. We refer throughout to imbalance on the unweighted specification since patterns of limited imbalance are generally shared irrespective of weighting schemes.

In Table A5 we report balance at the district level. In the unweighted specification, 0 (0) out of 18 covariates are imbalanced at the 5% (10%) level. In Table A6 we report balance in the endline survey sample when we assign district-level covariates to respondents. In the unweighted specification we find that 2 (2) out of 18 covariates are imbalanced. In Table A7 we report balance in the endline survey sample using individual-level covariates. In the unweighted specification, 1 (1) out of 4 covariates are imbalanced. In Table A8 we report balance at the polling station level. We find evidence of imbalance on 0 (1) covariates out of 18 covariates.

In Table A10 we assess evidence of imbalance on incumbent quality by treatment assignment. In the unweighted specification, we find imbalance on 0 (0) covariates out of 3. In Table A11 we test for imbalance at the candidate-level using our survey of candidates who ran in the election. Importantly, in Column 1 we demonstrate balanced response rates to our post-election survey across all candidates, incumbents and challengers across treatment groups. Using the full sample of candidates, we find imbalance on 1 (2) covariates out of 8. Restricting to incumbents we find imbalance on 0 (1) covariates. Restricting to predicted challengers we find imbalance on 2 (3) covariates. We consider balance at the candidate-level to be good particularly given our primarily descriptive employment of this data.

Overall we find little evidence of aggregate imbalance—whether on political or non-political

variables—and, when applicable, we control for the variables we blocked on throughout the analysis to deal with whatever imbalance that might exist.

A.3 Divergences from Pre-Analysis Plan

This study was pre-registered with EGAP (ID: 20171024AA) and AEA (ID: AEARCTR-0002553) under the title "Turning Up, Tuning In, Turning Out: Experimental Evidence from Liberia." Pre-registration took place before endline data collection and any data analysis. In this section we describe the differences between our PAP and the final paper, as well as the logic behind them.

A.3.1 Data and estimation

We reorganized some of categorizations of variables from the PAP to fit into more coherent groupings. This comprised combining 'Knowledge about candidates' and 'Beliefs over candidate competence' into 'Effects on beliefs about candidates;' and 'Voter coordination' and relevant parts of 'Debate exposure' into 'Effects on political engagement'.

As we discuss in the paper, we cross-randomized a separate intervention to vary the intensity of debate rebroadcasting (with either two or ten rebroadcasts) which ultimately had little effect. Since this additional intervention had no effect and candidates were unaware of rebroadcasting plans, we pool over rebroadcasting intensity for clarity of exposition. Importantly, however, we made no multiplicative hypotheses – rather, all our hypotheses were with respect to the overall *intensity* of the debates initiative and focused on those districts assigned to both high invitation intensity and high rebroadcasting intensity. We can demonstrate that all our key results also hold under this factorial design, but pooling rebroadcasting loses relatively little granularity and gains substantially in power. In Table A25 we show that the rebroadcasting intervention did not lead to significantly positive effects on debate exposure in either the full sample or the sample restricted to respondents in intensive-invitation districts, in spite of being correctly implemented. As discussed, this lack of effects stems from number of citizens who heard their district debates when they were only broadcast a few times.

The estimating equation we use in the paper is closest to what we called our 'base specification' in our PAP (Equation 5). We additionally pre-registered the possibility of constructing an individual-level instrument for the debate attendance of candidates, leveraging random assignment of candidates to debates with the incumbent and at different times of day in districts where more than one debate was held. We found such an instrument to be underpowered

due to the number of districts which only ended up holding one debate and so do not report results using it. We also pre-registered a local regression discontinuity design (Equation 8) leveraging quasi-random assignment to respondents being interviewed before or after the live debate in their district at baseline, but lacked sufficient within-district variation to pursue this. Finally, we pre-registered the use of one-tailed tests but report two-tailed tests throughout to be conservative.

We did not pre-register outcomes relating to 'on-the-ground' and 'radio' campaigning by candidates. Additionally, in our PAP we pre-registered the use of a jackknife measure of debate performance; given that actual debate performance is a function of treatment assignment we instead use the predicted measure of performance described in Section 4.2.

While our pre-registered hypotheses make reference to the distinction between incumbents and challenger candidates, particularly differential treatment effects of the intervention on their participation, we did not pre-register the descriptive analysis we perform regarding their intervention compliance behavior.

A.3.2 Hypotheses

We reorganised and grouped many of our pre-registered hypotheses, which were generally made with reference to individual outcome variables, into more coherent aggregated clusters. Out of the 27 hypotheses we pre-registered, results directly testing 19 of them are presented in the final paper.[32] The eight missing hypotheses fall into two categories. First, we do not report results relating to the hypotheses using within-district variation in whether citizens at baseline were interviewed before or after their district debate had been broadcast for the first time due to the lack of variation mentioned above.

Second, we do not report results for our pre-registered set of hypotheses relating to citizen attitudes towards the media and the electoral process. We anticipated that citizens in districts assigned to more intensive debates would update positively about the neutrality and contribution of the media to the electoral process due to the novelty and unbiasedness of the debate structure. We found little systematic evidence of this happening – we interpret this to be consistent with the campaigning response of incumbent candidates campaigning more aggressively on the radio in these districts where they were more likely to participate. We report these results for

[32] Broadly we aggregated hypotheses from 'Debate exposure and knowledge about candidates' and 'Beliefs about candidates' into the results on voter response; hypotheses from 'Preferences and voting behavior' into the results on voting outcomes; 'Media consumption, attitudes, and institutions' and 'Debate exposure and knowledge about candidates' into the results on voter response.

completeness in Table A24. The final paper therefore contains substantially less emphasis on the intervention affecting media credibility than our PAP.

A.4 Variable construction

In this appendix we document the construction of all variables used in the analysis. Unless otherwise noted, these variables come from our panel survey of citizens where we refer to specific items in our baseline and endline survey instruments using the format *wave-question*, where *wave* is represented by *B* (baseline) or *E* (endline) and *question* is simply the question on the relevant instrument. Both survey instruments can be found online at http://egap.org/registration/2899.

As described in Section 5, whenever we asked the same question in both baseline and endline we use the difference as an outcome. We preserve whether variables are discrete or continuous. For indices, we standardize each component such that units in the control group have zero mean and standard deviation of one then average them, following Kling et al. (2007).

Table 5:

- *Share of candidates*: share of the total candidates in that district who participated in their district debate.
- *Incumbent*: indicator for whether incumbent participated in their district debate.
- *Share of challengers*: share of the predicted challengers in that district (see Appendix A.1) who participated in their district debate.
- *Share of other candidates*: share of non-predicted candidates in that district (see Appendix A.1) who participated in their district debate.

Table 6:

- Panel 1:
  - *Main effect*: indicator for whether a respondent named a specific predicted leading candidate as their vote choice at endline (E-Q45).
  - *Interaction: Performance*: measure of predicted debate performance of a specific predicted leading candidate, generated by flexibly estimating observed measure of debate performance and predicting out-of-sample on candidates who did not participate.

  - *Interaction: Policy alignment*: measure of preference alignment between respondent and a specific predicted leading candidate. Defined as the share of the three priority issues the respondents name in their districts at baseline (B-Q13) that are shared with the priorities of a given candidate based on aggregating citizen perceptions of that candidates' priorities at baseline (B-Q15, B-Q17, B-Q19).

- Panel 2:
  - *Main effect*: Vote share of candidate at polling station-level.
  - *Interaction: Performance*: District-level analogue of respondent-level interaction.
  - *Interaction: Policy alignment*: District-level analogue of respondent-level interaction.

Table 8:

- *Debate listening index*: standardized index of:
  - *Change in heard debate*: indicator variable for whether the respondent heard their district debate between baseline (B-Q7) and endline (E-Q14) surveys.
  - *Heard debate*: indicator variable for whether the respondent heard their district debate at endline (E-Q14).
  - *Number of times heard*: continuous variable for the number of times respondents reported hearing their district debate at endline (E-Q15).
- *Debate knowledge index*: standardized index of:
  - *Debate winner attended debate:* indicator for whether respondent's named debate winner actually attended the debate (E-Q17).
  - *Stated share of participating candidates:* continuous variable for the share of candidates in respondent's district they recall participating in debate (E-Q16).
  - *Stated share of participating leading candidate:* continuous variable for the share of leading candidates in respondent's district they recall participating in debate (E-Q39.2, E-Q39.4, E-Q39.6).
- *Policy knowledge index*: standardized index of:
  - *Manager of CSDF*: change in whether respondents switch towards correctly identifying Representatives as the primary controller of CSDF (B-Q9, E-Q9).

  - *CSDF reporting requirement*: change in whether respondents correctly switch towards believing that it is a legal obligation to report CSDF allocation (B-Q10, E-Q10).
  - *CSDF citizen involvement requirement*: change in whether respondents correctly switch towards believing that it is a legal obligation to involve citizens in CSDF allocation decisions (B-Q11, E-Q11).

- *Political information demand index*: standardized index of:
  - *Change in radio listening*: Respondents' change in listening to radio between baseline (B-Q6, E-Q6).
  - *Demand for non-radio information sources*: How frequently respondents sought political information from non-radio sources such as newspapers, television and the internet (E-Q7).
  - *Change in political discussion with friends*: Change in how frequently respondents discussed political issues with friends, family, neighbors and other members of the community (B-Q8, E-Q8).

Table 9:

- *Certainty about competence*: standardized change in how sure respondents were about the competence of specific predicted leading candidates between baseline (B-Q22, B-Q24, B-Q26) and endline (E-Q34, E-Q36, E-Q38).
- *Certainty about issues*: standardized change in how sure respondents were about the priority issues of specific predicted leading candidates between baseline (B-Q16, B-Q18, B-Q20) and endline (E-Q28, E-Q30, E-Q32).
- *Beliefs about competence*: standardized change in how competent respondents believe specific predicted leading candidates were between baseline (B-Q21, B-Q23, B-Q25) and endline (E-Q33, E-Q35, E-Q37).
- *Learning about policy*: standardized change in the share of candidate priority issues that citizens name between baseline (B-Q15, B-Q17, B-Q19) and endline (E-27, E-29, E-31). We define candidate priorities using the aggregate of citizen beliefs over a given candidate's priorities measured in the baseline survey.

Table 10:

- *Ground*: standardized index of:

  - Whether candidate distributed leaflets or posters in respondent's community (E-Q41.1, E-Q41.2, E-Q41.3).
  - Whether candidate made campaign visits to respondent's community (E-Q41.1, E-Q41.2, E-Q41.3).
  - How frequently other people in their community voted for a given candidate in exchange for money, food or other gifts (E-Q40.1, E-Q40.2, E-Q40.3).

- *Radio*: standardized measure how how frequently respondents heard candidates on the radio in the two weeks before the election (E-Q39.1, E-Q39.3, E-Q39.5).

A.5 Effects on political engagement

The qualitative evidence we presented in Section 3.2 suggests that the debates had a catalyzing effect on political engagement. In this section we quantitatively test whether, by increasing the attendance of leading candidates, the invitation intervention affected the political engagement of citizens and ultimately their turnout.

In line with this, in Panel A there are strong treatment effects on a standardized coordination index capturing whether citizens discussed the debates with others and whether they believed these discussions led them to coordinate their vote choices. Specifically, citizens exhibit a 0.15 sd increase in coordination in treated districts. Both panels then provide strong evidence that, when exposed to debates in which relevant leading candidates participated, citizens demanded more political information and coordinated their vote choices.

Next, we provide evidence that the citizen engagement caused by the invitation intervention ultimately led to higher turnout. In Panel B of Table A1, we use administrative polling station data on turnout. Turnout in the 2017 House of Representatives election was on average 1.5 pp higher at polling stations in treated districts than the control mean of 70.1 percent (2%). Overall, the results provide strong evidence that greater candidate participation in the debates increased political engagement of citizens, which in turn generated actual electoral mobilization.

A1. Political Engagement

	(1)	(2)	(3)
A. Debate coordination index			
Invite	0.153**	0.158**	0.193***
	(0.065)	(0.061)	(0.061)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
B. Polling station-level turnout			
Invite	0.015***	0.015***	0.012**
	(0.005)	(0.005)	(0.005)
Control Mean	0.70	0.70	0.70
Observations	5383	5383	5383
Weight	No	Reg	1/PS

Outcome variables: Panel A: index of variables measuring (1) how much respondents discussed the debate with friends (2) how much this discussion led to coordinating their vote choices. Panel B: polling station-level turnout in House of Representatives election.

Specifications using citizen survey outcomes are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Specifications using polling station-level data are estimated using OLS and include block FE, district-level and polling station-level controls. For weighted specifications, 'PS' is the number of polling stations in that district and 'Reg' is the number of registered voters at that polling station. Standard errors clustered at the district-level in parentheses.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.6 Tables

A.6.1 Descriptive statistics

A2. Descriptive Statistics

	Mean	SD	Min	Max
A. District-level variables (n = 73)				
Scheduled debate week	4.18	1.39	1.00	8.00
Number of debates in district	2.08	0.66	1.00	4.00
Number of candidates (2017)	13.55	4.81	3.00	28.00
Incumbent ran in election (2017)	0.84	0.37	0.00	1.00
Share of repeat candidates (2017)	0.26	0.11	0.06	0.53
Log registered voters (2017)	10.23	0.40	9.27	11.06
1st voteshare (2011)	0.31	0.13	0.12	0.82
2nd voteshare (2011)	0.18	0.05	0.10	0.36
3rd voteshare (2011)	0.13	0.03	0.05	0.25
Voteshare HHI (2011)	0.19	0.11	0.07	0.69
Turnout (2011)	0.66	0.05	0.56	0.75
Log population density (2008)	−9.51	1.76	−11.91	−5.21
Share over 18 (2008)	0.48	0.02	0.43	0.54
Share with secondary education (2008)	0.15	0.05	0.04	0.28
Share with GSM coverage (2015)	0.71	0.30	0.01	1.00
Share owns a radio (2016)	0.74	0.12	0.38	1.00
Share gets radio news often (2016)	0.76	0.12	0.50	1.00
Avg. N radio stations covering each town (2016)	10.98	7.60	0.00	23.36
B. Individual-level variables (n = 4060)				
Male	0.75	0.43	0.00	1.00
Age	31.73	9.27	18.00	99.00
Highest education: primary school	0.07	0.26	0.00	1.00
Highest education: secondary school	0.29	0.46	0.00	1.00
Highest education: university	0.14	0.34	0.00	1.00
C. Polling station-level variables (n = 5386)				
Number of registered voters in PS (2017)	405.12	74.87	12.00	505.00
Number of PS in VRC	3.57	1.89	1.00	9.00
VRC added in 2017	0.10	0.31	0.00	1.00
Number of registered voters in VRC (2011)	1422.96	770.45	45.00	3995.00
Number of PS in VRC (2011)	3.36	1.59	1.00	9.00
Turnout (2011)	0.63	0.09	0.14	1.01
Share of invalid votes (2011)	0.07	0.03	0.00	0.35
PS covered by partner radio station	0.85	0.36	0.00	1.00
1/N radio stations covering PS	0.05	0.03	0.00	0.11
Urban PS	0.33	0.47	0.00	1.00
C. Polling station-level variables (n = 5386)				
Number of registered voters in PS (2017)	405.13	74.87	12.00	505.00
Number of PS in VRC	3.57	1.88	1.00	9.00
VRC added in 2017	0.10	0.31	0.00	1.00
Number of registered voters in VRC (2011)	1423.04	770.40	45.00	3995.00
Number of PS in VRC (2011)	3.36	1.59	1.00	9.00
Turnout (2011)	0.63	0.09	0.14	1.01
Share of invalid votes (2011)	0.07	0.03	0.00	0.35
PS covered by partner radio station	0.85	0.36	0.00	1.00
1/N radio stations covering PS	0.05	0.03	0.00	0.11
Urban PS	0.33	0.47	0.00	1.00

Table presents descriptive statistics relating to all control variables used in the regression analyses. *Sources:* **District-level variables:** Debate variables from Internews. All 2017 and 2011 variables come from National Elections Commission (NEC). All 2008 variables come from 2008 Population and Housing Census. 'Share with GSM coverage' comes from Collins Mobile Coverage Explorer. 'Share owns a radio' and 'Share gets radio news often' come from Afrobarometer. 'Avg. N radio stations covering each town' comes from Internews. **Individual-level variables:** All come from researchers' panel survey. **Polling station-level variables:** Radio station variables come from Internews. All other variables come from NEC. 90% of polling stations were in locations where a polling place (Voter Registration Center, VRC) existed in 2011. For new polling stations we assign district-level averages.

A3. Descriptive Statistics (Respondent level outcomes)

	Mean	SD	Min	Max
Table 5				
Share of candidates attended debate	0.59	0.22	0.11	1.00
Incumbent attended debate	0.35	0.48	0.00	1.00
Share of challengers attended debate	0.60	0.37	0.00	1.00
Table 8				
Heard debate between baseline and endline	0.09	0.29	0.00	1.00
Heard debate at endline	0.21	0.41	0.00	1.00
Number of times heard debate	0.46	1.06	0.00	24.00
Debate winner attended debate	0.32	0.47	0.00	1.00
Stated share of participating candidates	0.12	0.29	0.00	2.53
Stated incumbent debate participation	0.16	0.34	0.00	1.00
Switches to correct CSDF controller	0.25	0.43	0.00	1.00
Switches to correct requirement for CSDF reporting	0.25	0.43	0.00	1.00
Switches to correct requirement for CSDF citizen engagement	0.25	0.43	0.00	1.00
Change in radio listening	0.26	2.17	-6.00	6.00
Demand for non-radio information sources	5.01	2.15	0.00	7.00
Change in political discussion with friends	-0.07	2.30	-6.00	6.00
Table A1				
Discussed debate with friends	1.01	1.61	0.00	5.00
Discussion led to coordination	1.06	1.76	0.00	5.00
Table A1 (PS-level)				
Turnout	0.70	0.11	0.00	1.00
Table A1 (PS-level)				
Turnout	0.70	0.11	0.00	1.00

Table presents descriptive statistics relating to all unstandardized, disaggregated outcomes variables used in the respondent-level regressions as outcome variables.

A4. Descriptive Statistics (Respondent-candidate level outcomes)

	Incumbent				Challengers			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Table 6								
Switches to voting for candidate	0.16	0.51	-1.00	1.00	0.09	0.37	-1.00	1.00
Measure of debate performance	0.27	0.18	0.00	0.70	0.12	0.15	0.00	0.83
Measure of policy match	0.42	0.27	0.00	1.00	0.42	0.28	0.00	1.00
Table 9								
Change in certainty of candidate competence	0.08	1.40	-4.00	4.00	0.36	1.53	-4.00	4.00
Change in certainty of candidate priority issues	-0.03	1.63	-4.00	4.00	-0.02	1.58	-4.00	4.00
Change in assessment of candidate competence	0.28	1.44	-4.00	4.00	0.12	1.23	-4.00	4.00
Change in share of candidate priorities named	0.02	0.41	-1.00	1.00	0.03	0.42	-1.00	1.00
Table 10								
Frequency of hearing candidate on radio	2.95	1.14	0.00	5.00	2.67	1.13	0.00	5.00
Candidate distributed leaflets in locality	0.93	0.25	0.00	1.00	0.90	0.30	0.00	1.00
Candidate visited locality	0.81	0.39	0.00	1.00	0.74	0.44	0.00	1.00
Frequency of candidate vote buying	2.28	1.39	1.00	5.00	2.10	1.28	1.00	5.00
Table 6 (PS-level)								
Vote share	0.26	0.22	0.00	0.99	0.11	0.16	0.00	0.99
Measure of debate performance	0.25	0.19	0.00	0.68	0.12	0.13	0.00	0.75
Measure of policy alignment	0.43	0.09	0.16	0.59	0.43	0.09	0.00	0.64
Table 6 (PS-level)								
Vote share	0.26	0.22	0.00	0.99	0.11	0.16	0.00	0.99
Measure of debate performance	0.25	0.19	0.00	0.68	0.12	0.13	0.00	0.75
Measure of policy alignment	0.43	0.09	0.16	0.59	0.43	0.09	0.00	0.64

Table presents descriptive statistics relating to all unstandardized, disaggregated outcomes variables used in the respondent-candidate level regressions as outcome variables.

A.6.2 Balance

A5. District-level Balance

	VS of 1st place (2011)		VS of 1st place (2011)		VS of 3rd place (2011)	
Invite	-0.011	-0.015	-0.005	-0.008	-0.000	-0.001
	(0.024)	(0.025)	(0.010)	(0.010)	(0.007)	(0.006)
Control Mean	0.322	0.309	0.178	0.183	0.129	0.129
	Turnout (2011)		**Share ran in 2011**		**VS HHI (2011)**	
Invite	-0.004	-0.004	-0.012	-0.019	-0.006	-0.009
	(0.009)	(0.008)	(0.024)	(0.024)	(0.020)	(0.020)
Control Mean	0.661	0.667	0.258	0.266	0.198	0.189
	N. candidates (2017)		**Incumbent ran (2017)**		**Number of debates**	
Invite	-0.509	-0.670	-0.019	-0.052	-0.083	-0.131
	(0.833)	(0.928)	(0.082)	(0.078)	(0.126)	(0.140)
Control Mean	13.634	14.780	0.831	0.886	2.115	2.256
	Log pop. dens. (2008)		**Log reg. voters (2017)**		**GSM coverage (2016)**	
Invite	0.254	0.160	0.056	0.060	0.032	0.005
	(0.365)	(0.370)	(0.061)	(0.055)	(0.062)	(0.055)
Control Mean	-9.847	-9.108	10.214	10.351	0.645	0.773
	Share owns radio (2015)		**Share radio news (2015)**		**Radio coverage (2016)**	
Invite	-0.019	-0.034	-0.017	-0.038	-0.109	0.024
	(0.025)	(0.022)	(0.026)	(0.025)	(1.347)	(1.384)
Control Mean	0.755	0.773	0.767	0.780	10.051	12.830
	Debate week		**Share sec. ed. (2008)**		**Share 18+ (2008)**	
Invite	0.028	-0.030	0.004	0.003	0.000	0.001
	(0.215)	(0.209)	(0.010)	(0.010)	(0.004)	(0.004)
Control Mean	3.980	3.979	0.137	0.154	0.483	0.487
Observations	73	73	73	73	73	73
Weight	None	Reg	None	Reg	None	Reg

Descriptions of all variables can be found in Table A2.

All specifications are estimated using OLS and include block FE. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A6. Individual-level Balance (district covariates)

	VS of 1st place (2011)			**VS of 1st place (2011)**			**VS of 3rd place (2011)**		
Invite	0.007	-0.009	-0.014	-0.011	-0.005	-0.008	-0.000	-0.000	-0.001
	(0.025)	(0.021)	(0.021)	(0.009)	(0.008)	(0.009)	(0.007)	(0.006)	(0.005)
Control Mean	0.298	0.308	0.308	0.185	0.182	0.183	0.129	0.128	0.129
	Turnout (2011)			**Share ran in 2011**			**VS HHI (2011)**		
Invite	0.001	-0.004	-0.004	-0.043**	-0.012	-0.019	0.007	-0.005	-0.007
	(0.007)	(0.007)	(0.007)	(0.019)	(0.021)	(0.020)	(0.021)	(0.017)	(0.017)
Control Mean	0.665	0.659	0.667	0.275	0.270	0.266	0.181	0.189	0.188
	N. candidates (2017)			**Incumbent ran (2017)**			**Number of debates**		
Invite	-0.642	-0.549	-0.710	-0.072	-0.024	-0.057	-0.117	-0.091	-0.139
	(0.757)	(0.709)	(0.795)	(0.074)	(0.070)	(0.067)	(0.108)	(0.108)	(0.120)
Control Mean	15.084	13.963	14.833	0.912	0.860	0.889	2.291	2.152	2.265
	Log pop. dens. (2008)			**Log reg. voters (2017)**			**GSM coverage (2016)**		
Invite	0.207	0.246	0.149	0.072	0.055	0.059	0.029	0.032	0.004
	(0.317)	(0.312)	(0.317)	(0.052)	(0.052)	(0.047)	(0.050)	(0.053)	(0.047)
Control Mean	-8.975	-9.611	-9.096	10.342	10.215	10.354	0.795	0.695	0.774
	Share owns radio (2015)			**Share radio news (2015)**			**Radio coverage (2016)**		
Invite	-0.041**	-0.020	-0.035*	-0.035	-0.017	-0.039*	-0.172	-0.138	-0.010
	(0.020)	(0.022)	(0.019)	(0.025)	(0.022)	(0.022)	(1.222)	(1.159)	(1.190)
Control Mean	0.771	0.750	0.774	0.778	0.767	0.780	13.552	11.275	12.895
	Debate week			**Share sec. ed. (2008)**			**Share 18+ (2008)**		
Invite	-0.035	0.031	-0.026	0.006	0.004	0.002	-0.001	0.000	0.001
	(0.193)	(0.184)	(0.179)	(0.008)	(0.009)	(0.009)	(0.004)	(0.004)	(0.003)
Control Mean	4.162	4.148	3.980	0.160	0.146	0.154	0.488	0.485	0.487
Observations	4061	4061	4061	4061	4061	4061	4061	4061	4061
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

Descriptions of all variables can be found in Table A2.

All specifications are estimated using OLS and include block FE. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A7. Individual-level balance (individual covariates)

	Survey date			**Education**		
Invite	-1.063**	-0.796	-1.006*	0.067	0.102	0.081
	(0.523)	(0.572)	(0.573)	(0.086)	(0.084)	(0.075)
Control Mean	71.801	71.153	71.422	6.586	6.447	6.534
	Age			**Male**		
Invite	0.163	0.359	0.225	0.011	0.025	0.021
	(0.439)	(0.464)	(0.437)	(0.015)	(0.016)	(0.016)
Control Mean	31.728	32.103	31.877	0.746	0.744	0.740
Observations	4061	4061	4061	4061	4061	4061
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

All specifications are estimated using OLS and include block FE. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A8. Polling Station-level Balance (district covariates)

	VS of 1st place (2011)			**VS of 1st place (2011)**			**VS of 3rd place (2011)**		
Invite	-0.015	-0.010	-0.015	-0.009	-0.005	-0.009	-0.001	-0.000	-0.001
	(0.021)	(0.021)	(0.021)	(0.009)	(0.008)	(0.009)	(0.005)	(0.006)	(0.005)
Control Mean	0.308	0.309	0.309	0.183	0.184	0.184	0.129	0.129	0.129
	Turnout (2011)			**Share ran in 2011**			**VS HHI (2011)**		
Invite	-0.004	-0.004	-0.004	-0.016	-0.012	-0.018	-0.008	-0.006	-0.008
	(0.007)	(0.007)	(0.007)	(0.020)	(0.021)	(0.020)	(0.017)	(0.017)	(0.018)
Control Mean	0.665	0.667	0.667	0.266	0.266	0.266	0.188	0.188	0.188
	N. candidates (2017)			**Incumbent ran (2017)**			**Number of debates**		
Invite	-0.648	-0.506	-0.677	-0.050	-0.019	-0.053	-0.124	-0.083	-0.130
	(0.771)	(0.715)	(0.798)	(0.067)	(0.070)	(0.066)	(0.118)	(0.109)	(0.121)
Control Mean	14.651	14.794	14.794	0.883	0.890	0.890	2.239	2.256	2.256
	Log pop. dens. (2008)			**Log reg. voters (2017)**			**GSM coverage (2016)**		
Invite	0.178	0.255	0.154	0.062	0.056	0.060	0.010	0.032	0.005
	(0.322)	(0.314)	(0.320)	(0.048)	(0.052)	(0.047)	(0.048)	(0.053)	(0.047)
Control Mean	-9.223	-9.097	-9.097	10.325	10.352	10.352	0.753	0.775	0.775
	Share owns radio (2015)			**Share radio news (2015)**			**Radio coverage (2016)**		
Invite	-0.031	-0.019	-0.033*	-0.037*	-0.017	-0.038*	-0.001	-0.100	-0.065
	(0.019)	(0.022)	(0.019)	(0.021)	(0.022)	(0.022)	(1.201)	(1.157)	(1.199)
Control Mean	0.770	0.773	0.773	0.778	0.780	0.780	12.485	12.919	12.919
	Debate week			**Share sec. ed. (2008)**			**Share 18+ (2008)**		
Invite	-0.041	0.029	-0.046	0.003	0.004	0.002	0.002	0.000	0.001
	(0.176)	(0.185)	(0.178)	(0.009)	(0.009)	(0.009)	(0.003)	(0.004)	(0.003)
Control Mean	4.007	3.987	3.987	0.152	0.154	0.154	0.486	0.487	0.487
Observations	5383	5383	5383	5383	5383	5383	5383	5383	5383
Weight	No	1/PS	Reg	No	1/PS	Reg	No	1/PS	Reg

Descriptions of all variables can be found in Table A2.

All specifications are estimated using OLS and include block FE. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A9. Polling Station-level Balance (PS covariates)

	N. PS in VRC (2017)			N. PS in VRC (2011)		
Invite	-0.099	0.008	-0.121	-0.114	-0.011	-0.142
	(0.247)	(0.226)	(0.249)	(0.185)	(0.183)	(0.186)
Control Mean	3.605	3.807	3.807	3.399	3.571	3.571
	New VRC			**Urban**		
Invite	-0.014	-0.008	-0.015	0.067	0.079	0.063
	(0.014)	(0.016)	(0.013)	(0.058)	(0.056)	(0.059)
Control Mean	0.111	0.101	0.101	0.299	0.326	0.326
	Reg. Voters (2017)			**Reg. Voters (2011)**		
Invite	2.247	4.668	0.686	-53.204	0.239	-67.706
	(6.170)	(6.416)	(4.775)	(90.770)	(89.758)	(90.943)
Control Mean	403.792	418.480	418.480	1439.977	1526.086	1526.086
	Turnout (2011)			**Invalid votes (2011)**		
Invite	0.014	0.011	0.015*	0.002	0.004	0.002
	(0.009)	(0.009)	(0.009)	(0.003)	(0.003)	(0.003)
Control Mean	0.620	0.620	0.620	0.066	0.066	0.066
	Radio covered			**Radio intensity**		
Invite	-0.016	0.012	-0.025	0.001	0.003	-0.000
	(0.054)	(0.052)	(0.054)	(0.004)	(0.004)	(0.004)
Control Mean	0.848	0.858	0.858	0.046	0.045	0.045
Observations	5383	5383	5383	5383	5383	5383
Weight	No	1/PS	Reg	No	1/PS	Reg

Descriptions of all variables can be found in Table A2. Analysis throughout is at the polling station-level, where multiple polling stations exist within a single location called a VRC. 1780/2080 VRCs existed in the 2011 election; for these 300 new VRCs we assign district-level averages.

All specifications are estimated using OLS and include block FE. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A10. Incumbent Balance

	Attendance		Absent		Distant	
	(1)	(2)	(3)	(4)	(5)	(6)
Invite	0.034	0.019	-0.003	0.013	-0.025	-0.023
	(0.032)	(0.030)	(0.020)	(0.024)	(0.028)	(0.024)
Mean	0.791	0.807	0.117	0.114	0.079	0.066
Observations	73	73	73	73	73	73
Weight	None	Reg	None	Reg	None	Reg

Outcome variables are plenary session attendance measures taken from legislator scorecards for 2016. Legislators either attend, are absent, or are away from Monrovia for each plenary session.

All specifications are estimated using OLS and include block FE. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A11. Candidate-level Balance

	Response (1)	Age (2)	Univ. ed. (3)	Radio (4)	Male (5)	Ran before (6)	Gov job (7)	Advocacy (8)	NGO (9)
A. All candidates									
Invite	-0.028	-0.379	0.012	0.020	0.009	-0.024	-0.028	0.064	0.065
	(0.031)	(1.064)	(0.049)	(0.013)	(0.027)	(0.032)	(0.030)	(0.025)	(0.036)
Control Mean	0.63	48.48	0.55	0.03	0.85	0.30	0.34	0.83	0.38
Observations	984	608	612	612	612	612	612	612	612
B. Incumbents									
Invite	0.139	2.394	0.061	-0.030	0.121	0.000	-0.394	-0.121	0.061
	(0.131)	(4.085)	(0.210)	(0.200)	(0.174)	(.)	(0.212)	(0.174)	(0.255)
Control Mean	0.42	55.31	0.69	0.15	0.77	1.00	0.69	0.92	0.23
Observations	64	31	31	31	31	31	31	31	31
C. Challengers									
Invite	0.010	3.875	-0.005	-0.002	0.027	0.134	0.036	0.155	0.175
	(0.083)	(1.655)	(0.106)	(0.052)	(0.067)	(0.102)	(0.086)	(0.065)	(0.103)
Control Mean	0.64	47.31	0.65	0.06	0.90	0.35	0.31	0.81	0.29
Observations	155	99	102	102	102	102	102	102	102

Panel A presents unweighted balance tests for the full set of candidates in survey, Panel B restricts to incumbent candidates, Panel C restricts to predicted challenger candidates. Outcome variables are: response rate to survey; age in years; indicator for whether candidate completed university; indicator for whether they own or manage a radio station; indicator for candidate being male; indicator for candidate having run for office before; indicator for candidate having a government job before; indicator for candidate having advocacy experience; indicator for candidate working for an NGO before.

All specifications are estimated using OLS and include block FE. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A12. Candidate-level Balance (weighted)

	Response (1)	Age (2)	Univ ed (3)	Radio (4)	Male (5)	Ran before (6)	Gov job (7)	Advocacy (8)	NGO (9)
A. All candidates									
Invite	-0.045	-0.171	-0.013	0.030	0.025	-0.031	-0.018	0.043	0.057
	(0.033)	(0.899)	(0.056)	(0.017)	(0.031)	(0.036)	(0.039)	(0.027)	(0.037)
Mean	0.63	48.48	0.55	0.03	0.85	0.30	0.34	0.83	0.38
Observations	984	608	612	612	612	612	612	612	612
B. Challengers									
Invite	-0.037	4.918	0.042	0.008	0.008	0.153	0.054	0.144	0.165
	(0.064)	(1.815)	(0.105)	(0.064)	(0.068)	(0.106)	(0.094)	(0.068)	(0.101)
Mean	0.78	47.31	0.65	0.06	0.90	0.35	0.31	0.81	0.29
Observations	155	99	102	102	102	102	102	102	102

In this table we weight observations by the inverse of the number of responding candidate types in a given district. Since there is one incumbent per district, for incumbents this would be identical to Panel B of Table A11 and thus we omit it. Panel A presents balance tests for the full set of candidates in survey, Panel B restricts predicted challenger candidates. Outcome variables are: response rate to survey; age in years; indicator for whether candidate completed university; indicator for whether they own or manage a radio station; indicator for candidate being male; indicator for candidate having run for office before; indicator for candidate having a government job before; indicator for candidate having advocacy experience; indicator for candidate working for an NGO before.

All specifications are estimated using OLS and include block FE. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A.6.3 Demography-weighted specifications

A13. Candidate Debate Participation (demography-weighted)

	(1)	(2)
A. Share of candidates		
Invite	0.077**	0.077**
	(0.034)	(0.033)
Control Mean	0.542	0.557
Observations	4060	4060
Weight	No	Dem
B. Incumbent		
Invite	0.212**	0.213**
	(0.083)	(0.082)
Control Mean	0.280	0.299
Observations	4060	4060
Weight	No	Dem
C. Share of challengers		
Invite	0.212***	0.214***
	(0.074)	(0.073)
Control Mean	0.492	0.528
Observations	4060	4060
Weight	No	Dem
D. Share of other candidates		
Invite	0.003	0.002
	(0.030)	(0.029)
Control Mean	0.562	0.575
Observations	3991	3991
Weight	No	Dem

Outcome variables are the share of the respective set of candidates (winner, actual challenger) who attended a debate out of all candidates in that district. Actual challengers are defined as candidates who ranked in the top three in their race in the election but were not the incumbent. Actual other candidates are those who did not rank in the top three. Panels A and B have 4060 observations; Panel C has fewer due to only three candidates running in two districts (and hence no 'actual other candidates' defined).

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Dem' weights observations to be representative at the district-level with respect to gender and education. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A14. Voting Outcomes (demography-weighted)

	Main effect		Interaction term: Std. Performance		Interaction term: Std. Policy alignment	
1. Respondent-level	(1)	(2)	(3)	(4)	(5)	(6)
A. Incumbent						
Invite	0.045**	0.042*	0.049**	0.046**	0.041**	0.038*
	(0.020)	(0.022)	(0.020)	(0.023)	(0.019)	(0.022)
Invite × Std. performance			0.067*	0.063		
			(0.038)	(0.040)		
Invite × Std. policy alignment					0.042***	0.051**
					(0.016)	(0.021)
Control Mean	0.278	0.278	0.278	0.278	0.278	0.278
Observations	3496	3496	3496	3496	3496	3496
B. Challengers						
Invite	-0.048***	-0.052***	-0.051***	-0.056***	-0.048***	-0.052***
	(0.014)	(0.016)	(0.017)	(0.019)	(0.014)	(0.016)
Invite × Std. performance			-0.015	-0.026		
			(0.046)	(0.051)		
Invite × Std. policy alignment					-0.006	-0.004
					(0.007)	(0.008)
Control Mean	0.156	0.160	0.156	0.160	0.156	0.160
Observations	8684	8684	8684	8684	8684	8684
Weight	No	Dem	No	Dem	No	Dem

Panel 1: Outcome variable is an indicator for whether respondent expressed voting for the leading candidate at endline. Columns 1-3 report the main effects, Columns 4-6 include interactions of treatment assignment with standardized candidate-level measures of predicted debate performance, and Columns 7-9 include interactions with standardized respondent-candidate-level measures of preference alignment measured at baseline. Section 4.2 explains these interaction terms further.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Dem' weights observations to be representative at the district-level with respect to gender and education. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A15. Debate Exposure (demography-weighted)

	(1)	(2)
A. Debate listening index		
Invite	0.304***	0.294**
	(0.102)	(0.113)
Observations	4060	4060
Weight	No	Dem
B. Debate knowledge index		
Invite	0.125**	0.120**
	(0.063)	(0.056)
Observations	4060	4060
Weight	No	Dem
C. Policy knowledge index		
Invite	0.156*	0.110
	(0.089)	(0.098)
Observations	4060	4060
Weight	No	Dem
D. Political information acquisition		
Invite	0.251***	0.230**
	(0.078)	(0.114)
Observations	4060	4060
Weight	No	Dem

Outcome variables: Panel A: a standardized index of (1) indicator for whether the respondent had not heard their district debate at baseline but had at endline (2) indicator for whether the respondent had heard the debate at endline (3) the number of times the respondent had heard the debate at endline. Panel B: a standardized index of (1) indicator for whether the respondent's stated debate winner actually attended the debate (2) share of candidates respondent claims participated (3) share of predicted leading candidates respondent claims participated. Panel C: a standardized index of the change in how many factual questions about CSDF management respondents answered correctly between baseline and endline. Panel D: a standardized index of (1) change in how much respondents listened to the radio (2) change in how much they discussed politics with their friends (3) how much they accessed other sources of political information.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Dem' weights observations to be representative at the district-level with respect to gender and education. Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A16. Updating About Candidates (demography-weighted)

	Certainty about competence		Certainty about policy	
1. Uncertainty	(1)	(2)	(3)	(4)
A. Incumbent				
Invite	0.178*	0.162	0.169**	0.211***
	(0.105)	(0.107)	(0.074)	(0.078)
Observations	3496	3496	3496	3496
Weight	No	Dem	No	Dem
B. Challengers				
Invite	0.037	0.047	0.139**	0.072
	(0.066)	(0.078)	(0.061)	(0.070)
Observations	8684	8684	8684	8684
Weight	No	Dem	No	Dem

	Beliefs about competence		Learning about policy	
2. Levels	(1)	(2)	(3)	(4)
A. Incumbent				
Invite	0.098	0.134	0.089	0.054
	(0.075)	(0.099)	(0.065)	(0.067)
Observations	3496	3496	3496	3496
Weight	No	Dem	No	Dem
B. Challengers				
Invite	-0.078	-0.055	0.038	0.012
	(0.075)	(0.083)	(0.065)	(0.084)
Observations	8684	8684	8684	8684
Weight	No	Dem	No	Dem

Panels 1.A and 1.B: the outcome variable in columns 1-3 is the standardized change in certainty respondents express about candidate competence between baseline and endline, and in columns 4-6 is the standardized change in certainty respondents express about candidate priority issues between baseline and endline. Panels 2.A and 2.B: the outcome variable in columns 1-3 is the standardized change in respondent perception about candidate competence between baseline and endline, and in columns 4-6 is the standardized learning that respondents reflect about candidate priority issues between baseline and endline.
All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Dem' weights observations to be representative at the district-level with respect to gender and education. Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A17. Candidate Campaigning (demography-weighted)

	Ground		Radio	
	(1)	(2)	(3)	(4)
A. Incumbent				
Invite	-0.055	-0.106**	0.082**	0.083**
	(0.043)	(0.044)	(0.037)	(0.038)
Observations	3492	3492	3496	3496
Weight	No	Dem	No	Dem
B. Challengers				
Invite	-0.060*	-0.089***	-0.025	0.003
	(0.031)	(0.032)	(0.028)	(0.032)
Observations	8676	8676	8684	8684
Weight	No	Dem	No	Dem

Outcome variable in columns 1-3 is a standardized index of how often candidates (1) visited (2) distributed leaflets (3) bought votes in respondents' communities during campaigning. Outcome variable in columns 4-6 is a standardized measure of how often respondents heard candidates on the radio in the two weeks before the election.
All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Dem' weights observations to be representative at the district-level with respect to gender and education. Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.6.4 Effects on individual outcomes

A18. Debate Exposure (supplementary)

	(1)	(2)	(3)
A. Debate listening index			
(1) Change in heard debate			
Invite	0.080***	0.080***	0.102***
	(0.029)	(0.026)	(0.027)
Control Mean	0.084	0.082	0.082
(2) Heard debate			
Invite	0.038*	0.035*	0.050**
	(0.022)	(0.020)	(0.021)
Control Mean	0.195	0.202	0.193
(3) Number of times heard			
Invite	0.085*	0.104**	0.120***
	(0.045)	(0.046)	(0.045)
Control Mean	0.420	0.440	0.420
B. Debate knowledge index			
(1) Debate winner attended debate			
Invite	0.075**	0.078***	0.096***
	(0.031)	(0.029)	(0.029)
Control Mean	0.291	0.297	0.283
(2) Stated share of participating candidates			
Invite	0.023	0.024	0.031**
	(0.015)	(0.015)	(0.014)
Control Mean	0.111	0.115	0.111
(3) Stated share of participating leading candidates			
Invite	0.030*	0.026*	0.039**
	(0.017)	(0.016)	(0.016)
Control Mean	0.145	0.153	0.148
C. Policy knowledge index			
(1) Manager of CSDF			
Invite	0.041	0.053	0.044
	(0.029)	(0.038)	(0.032)
Control Mean	0.237	0.231	0.243
(2) CSDF reporting requirement			
Invite	-0.011	0.018	-0.008
	(0.032)	(0.035)	(0.033)
Control Mean	0.247	0.249	0.247
(3) CSDF citizen involvement requirement			
Invite	0.094**	0.113**	0.114**
	(0.040)	(0.051)	(0.045)
Control Mean	0.246	0.246	0.244
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
D. Political information demand index			
(1) Change in radio listening			
Invite	0.285**	0.451***	0.370**
	(0.134)	(0.164)	(0.160)
Control Mean	0.233	0.213	0.232
(2) Demand for non-radio information sources			
Invite	0.143*	0.168*	0.179**
	(0.075)	(0.085)	(0.083)
Control Mean	4.970	4.800	4.932
(3) Change in political discussion with friends			
Invite	0.510***	0.516**	0.491**
	(0.189)	(0.226)	(0.193)
Control Mean	-0.124	-0.106	-0.117

All outcome variables are described in Section A.4.
All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A19. Candidate Campaigning (supplementary)

	(1)	(2)	(3)
A. Incumbent			
(1) Candidate visited community			
Invite	-0.041**	-0.050**	-0.048**
	(0.016)	(0.023)	(0.021)
Control Mean	0.823	0.824	0.820
(2) Candidate distributed leaflets			
Invite	-0.014	-0.013	-0.014
	(0.009)	(0.013)	(0.011)
Control Mean	0.935	0.933	0.932
(3) Vote buying for candidate			
Invite	0.008	0.056	0.073
	(0.053)	(0.052)	(0.050)
Control Mean	2.258	2.223	2.231
Observations	3493	3493	3493
B. Challenger			
(1) Candidate visited community			
Invite	-0.036***	-0.029**	-0.036***
	(0.012)	(0.011)	(0.012)
Control Mean	0.756	0.753	0.753
(2) Candidate distributed leaflets			
Invite	-0.021***	-0.019**	-0.022**
	(0.007)	(0.008)	(0.008)
Control Mean	0.905	0.908	0.907
(3) Vote buying for candidate			
Invite	-0.014	-0.030	-0.007
	(0.037)	(0.041)	(0.040)
Control Mean	2.109	2.145	2.132
Observations	8678	8678	8678
Weight	No	1/Obs	Reg/Obs

All outcome variables are described in Section A.4.
All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.6.5 Other tables

A20. Candidate Debate Participation (district level)

	(1)	(2)
A. Share of candidates		
Invite	0.087	0.111*
	(0.056)	(0.056)
Control Mean	0.572	0.557
Observations	73	73
Weight	No	Reg
B. Incumbent		
Invite	0.275**	0.312**
	(0.127)	(0.127)
Control Mean	0.371	0.300
Observations	73	73
Weight	No	Reg
C. Share of challengers		
Invite	0.164	0.233*
	(0.122)	(0.117)
Control Mean	0.552	0.528
Observations	73	73
Weight	No	Reg
D. Share of other candidates		
Invite	0.014	0.011
	(0.055)	(0.056)
Control Mean	0.583	0.575
Observations	71	71
Weight	No	Reg

Outcome variables are the share of the respective set of candidates (all, incumbent, predicted challenger, other candidate) who attended a debate out of all candidates in that district. Panels A-C have 73 observations; Panel D has fewer due to only three candidates running in two districts (and hence no 'other candidates' defined). For weighted specifications, 'Reg' is the number of registered voters in that district.

All specifications are estimated using OLS and include block FE. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A21. Candidate Debate Participation (supplementary)

	(1)	(2)	
A. Election winner			
Invite	0.253**	0.202**	0.275***
	(0.097)	(0.089)	(0.093)
Control Mean	0.501	0.520	0.474
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
B. Share of actual challengers			
Invite	0.267***	0.176**	0.237***
	(0.068)	(0.068)	(0.068)
Control Mean	0.488	0.572	0.525
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
C. Share of actual other candidates			
Invite	0.018	0.029	0.036
	(0.039)	(0.037)	(0.038)
Control Mean	0.563	0.584	0.584
Observations	3991	3991	3991
Weight	No	1/Obs	Reg/Obs

Outcome variables are the share of the respective set of candidates (winner, actual challenger) who attended a debate out of all candidates in that district. Actual challengers are defined as candidates who ranked in the top three in their race in the election but were not the incumbent. Actual other candidates are those who did not rank in the top three. Panels A and B have 4060 observations; Panel C has fewer due to only three candidates running in two districts (and hence no 'actual other candidates' defined).

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A22. Debate Rebroadcasting

	(1)	(2)	(3)
A. Radio monitors			
Invite	0.387	-0.700	-0.206
	(0.998)	(0.860)	(0.974)
Control Mean	5.230	5.618	5.466
B. Radio survey			
Invite	1.055	0.326	0.746
	(1.044)	(0.906)	(0.954)
Control Mean	7.473	7.702	7.698
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs

Outcome variables are, in Panel A, the number of contracted rebroadcasts confirmed by radio monitors and, in panel B, Number of rebroadcasts based on survey of radio stations, including those not contracted to rebroadcast but recorded as being present in the debate venue.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A23. Vote Choice Reason

	Campaign promises			**Expectations**		
	(1)	(2)	(3)	(4)	(5)	(6)
Invite	0.006	0.002	0.000	0.018	0.005	0.024
	(0.010)	(0.012)	(0.012)	(0.019)	(0.021)	(0.018)
Control Mean	0.033	0.036	0.035	0.067	0.064	0.063
	Competence			**Experience**		
	(1)	(2)	(3)	(4)	(5)	(6)
Invite	0.013	0.016	0.001	-0.024	-0.001	-0.008
	(0.030)	(0.032)	(0.033)	(0.026)	(0.030)	(0.028)
Control Mean	0.196	0.209	0.201	0.262	0.254	0.258
Observations	4060	4060	4060	4060	4060	4060
Weight	No	1/Obs	Reg/Obs	No	1/Obs	Reg/Obs

Outcome variable is an indicator for whether respondents switched towards citing candidate campaign promises, expectated policy by the candidate, candidate competence or candidate experience as their main reason for their vote choice.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Standard errors clustered at the district-level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A24. Broader Consequences

	(1)	(2)	(3)
A. Media: trust and bias			
Invite	-0.007	-0.003	-0.004
	(0.013)	(0.013)	(0.013)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
B. Media: helps democracy			
Invite	-0.003	-0.004	-0.015
	(0.020)	(0.026)	(0.025)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
C. Electoral attitudes			
Invite	0.010	0.013	0.004
	(0.027)	(0.031)	(0.029)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs

Outcome variables are all z-score indices. Panel A: extent to which the media (1) was unbiased during election (2) gave equal coverage of candidates (3) is trustworthy. Panel B: media (1) helps select competent representatives (2) ensures representatives reflect views of voters. Panel C: elections (1) help select competent representatives (2) ensure representatives reflect views of voters.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A25. Rebroadcasting Intervention

	(1)	(2)	(3)
A. Full sample			
Rebroadcast	0.177	0.139	0.178
	(0.112)	(0.108)	(0.126)
Observations	4060	4060	4060
Weight	No	1/Obs	Reg/Obs
	(1)	(2)	(3)
B. Respondents in intensive-invite districts			
Rebroadcast	-0.025	-0.026	0.026
	(0.186)	(0.157)	(0.191)
Observations	2252	2252	2252
Weight	No	1/Obs	Reg/Obs

Outcome variable is our standardized index of debate exposure. In Panel A, we show no overall effects on debate exposure. In Panel B, we show no effects on debate exposure if we restrict to respondents in those districts assigned to high invitation intensity.

All specifications are estimated using OLS and include block FE, enumerator FE, district-level and individual-level controls. For weighted specifications, 'Obs' is the number of observations in that district and 'Reg' is the number of registered voters in that district. Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.